

UNITED STATES AIR FORCE
SUMMER RESEARCH PROGRAM -- 1998
SUMMER FACULTY RESEARCH PROGRAM FINAL REPORTS

VOLUME 3
PHILLIPS LABORATORY

RESEARCH & DEVELOPMENT LABORATORIES
5800 Uplander Way
Culver City, CA 90230-6608

Program Director, RDL
Gary Moore

Program Manager, AFOSR
Colonel Jan Cervený

Program Manager, RDL
Scott Licoscó

Program Administrator, RDL
Johnetta Thompson

Program Administrator, RDL
Rebecca Kelly-Clemmons

Submitted to:

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
Bolling Air Force Base
Washington, D.C.
December 1998

20010319 045

AQMOI-06-1173

PREFACE

Reports in this volume are numbered consecutively beginning with number 1. Each report is paginated with the report number followed by consecutive page numbers, e.g., 1-1, 1-2, 1-3; 2-1, 2-2, 2-3.

This document is one of a set of 15 volumes describing the 1998 AFOSR Summer Research Program. The following volumes comprise the set:

<u>VOLUME</u>	<u>TITLE</u>
1	Program Management Report
	<i>Summer Faculty Research Program (SFRP) Reports</i>
2	Armstrong Laboratory
3	Phillips Laboratory
4	Rome Laboratory
5A & 5B	Wright Laboratory
6	Arnold Engineering Development Center, Air Logistics Centers, United States Air Force Academy and Wilford Hall Medical Center
	<i>Graduate Student Research Program (GSRP) Reports</i>
7	Armstrong Laboratory
8	Phillips Laboratory
9	Rome Laboratory
10	Wright Laboratory
11	Arnold Engineering Development Center, and Wilford Hall Medical Center
	<i>High School Apprenticeship Program (HSAP) Reports</i>
12	Armstrong Laboratory
13	Phillips Laboratory
14	Rome Laboratory
15A, 15B & 15C	Wright Laboratory

REPORT DOCUMENTATION PAGE			Approved 0704-0188	
<small>Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing existing information, gathering the data, reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Project Director, Washington, DC 20503.</small>				
1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE December, 1998		
4. TITLE AND SUBTITLE 1998 Summer Research Program (SRP), Summer Faculty Research Program (SFRP), Final Reports, Volume 3, Phillips Laboratory		5. FUNDING NUMBERS F49620-93-C-0063		
6. AUTHOR(S) Gary Moore				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Research & Development Laboratories (RDL) 5800 Uplander Way Culver City, CA 90230-6608		8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Office of Scientific Research (AFOSR) 801 N. Randolph St. Arlington, VA 22203-1977		10. SPONSORING/MONITORING AGENCY REPORT NUMBER		
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION AVAILABILITY STATEMENT Approved for Public Release			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) The United States Air Force Summer Research Program (USAF-SRP) is designed to introduce university, college, and technical institute faculty members, graduate students, and high school students to Air Force research. This is accomplished by the faculty members (Summer Faculty Research Program, (SFRP)), graduate students (Graduate Student Research Program (GSRP)), and high school students (High School Apprenticeship Program (HSAP)) being selected on a nationally advertised competitive basis during the summer intersession period to perform research at Air Force Research Laboratory (AFRL) Technical Directorates, Air Force Air Logistics Centers (ALC), and other AF Laboratories. This volume consists of a program overview, program management statistics, and the final technical reports from the SFRP participants at the Phillips Laboratory.				
14. SUBJECT TERMS Air Force Research, Air Force, Engineering, Laboratories, Reports, Summer, Universities, Faculty, Graduate Student, High School Student			15. NUMBER OF PAGES	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UL	

AFRL-SR-BL-TR-00-

0778

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to **stay within the lines** to meet **optical scanning requirements**.

Block 1. Agency Use Only (*Leave blank*).

Block 2. Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C - Contract
G - Grant
PE - Program
Element

PR - Project
TA - Task
WU - Work Unit
Accession No.

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es). Self-explanatory.

Block 10. Sponsoring/Monitoring Agency Report Number. *//if known*

Block 11. Supplementary Notes. Enter information not included elsewhere such as: Prepared in cooperation with....; Trans. of....; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

DOD - See DoDD 5230.24, "Distribution Statements on Technical Documents."

DOE - See authorities.

NASA - See Handbook NHB 2200.2.

NTIS - Leave blank.

Block 12b. Distribution Code.

DOD - Leave blank.

DOE - Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.
Leave blank.

NASA - Leave blank.

NTIS -

Block 13. Abstract. Include a brief (*Maximum 200 words*) factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code (*NTIS only*).

Blocks 17. - 19. Security Classifications. Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.

SFRP FINAL REPORT TABLE OF CONTENTS

i-x

1. INTRODUCTION	1
2. PARTICIPATION IN THE SUMMER RESEARCH PROGRAM	2
3. RECRUITING AND SELECTION	3
4. SITE VISITS	4
5. HBCU/MI PARTICIPATION	4
6. SRP FUNDING SOURCES	5
7. COMPENSATION FOR PARTICIPATIONS	5
8. CONTENTS OF THE 1996 REPORT	6

APPENDICIES:

A. PROGRAM STATISTICAL SUMMARY	A-1
B. SRP EVALUATION RESPONSES	B-1

SFRP FINAL REPORTS

SRP Final Report Table of Contents

Author	University/Institution Report Title	Armstrong Laboratory Directorate	Vol-Page
DR Harvey Babkoff	Ramat-Gan , Rechovoth Israel , The Impact Of Bright Light and a Moderate Caffeine Dose on Nocturnal Performance: A Preliminary Exp	AFRL/HEP _____	2- 1
DR Michael P Dooley	Iowa State University , Ames , IA Exposure of Female Rats to a 35GHz Electromagnetic Field on Day 1 of Gestation does not Alter Pregna	AFRL/HED _____	2- 2
DR Brent D Foy	Wright State University , Dayton , OH Kinetic Modeling of Slow Dissociation of Bromosulphophthalein from Albumin in Perfusedd Rat Liver:	AFRL/HES _____	2- 3
DR Kenneth A Graetz	University of Dayton , Dayton , OH Negotiation at a Distance: Why You Might Want to Use the Telephone	AFRL/HEN _____	2- 4
DR Verlin B Hinsz	North Dakota State University , Fargo , ND Conceptualizing Crew Performance in Dynamic Operational Environments A Hierachy of Embedded Acton-Co	AFRL/HEC _____	2- 5
DR Nandini Kannan	Univ of Texas at San Antonio , San Antonio , TX Statistical Models for Altitude Decompression Sickness	AFRL/HEP _____	2- 6
DR Ramaswamy Ramesh	Research Foundation of SUNY , Buffalo , NY Aircraft and DT: Modeling and Analysis of Training Effectiveness, Flight Tradeoffs, Costs and Resour	AFRL/HEA _____	2- 7
DR Nancy J Stone	Creighton University , Omaha , NE Reliability and Validity Testing of the Student Characteristics Scale	AFRL/HEJ _____	2- 8
DR Robin D Thomas	Miami University , Oxford , OH Preliminary Decisiion Analysis of the Data Exploitation,mission Planning and Communication (DEMPC) S	AFRL/HEJ _____	2- 9
DR Ram C Tripathy	Univ of Texas at San Antonio , San Antonio , TX The effect of Repeated Measurements of the Variance of The Estiamted od the ahalf Life of Dioxin in	AFRL/HED _____	2- 10
DR Edward W Wolfe	University of Florida , Gainesville , FL Detecting bidimensionality in Response Data: An Empirical Task Analysis Technique	AFRL/HEJ _____	2- 11

SRP Final Report Table of Contents

Author	University/Institution Report Title	Phillips Laboratory Directorate	Vol-Page
DR Graham R Allan	National Avenue , Las Vegas , NM Temporal Characterisation of a Synchronously-Pumped Periodically-Poled Lithium Niobate Optical Param	AFRL/DEL _____	3- 1
DR Mark J Balas	Univ of Colorado at Boulder , Boulder , CO Stable Controller Design for Deployable Precision Structures Using Perturbation Theory	AFRL/VSD _____	3- 2
DR Neb Duric	University of New Mexico , Albuquerque , NM Image Recovery Using Phase Diversity	AFRL/DEB _____	3- 3
DR Arthur H Edwards	University of N. C.- Charlotte , Charlotte , NC Theory of Hydrogen In SiO2	AFRL/VSS _____	3- 4
DR Claudio O Egalon	University of Puerto Rico , Mayaguez , PR Investigating The use of Optical Fiber as Optical Delay Line For Adaptive Optics Systems	AFRL/DEB _____	3- 5
DR Jeffrey F Friedman	University of Puerto Rico , San Juan , PR Low Light Level Adaptive Optics Applied to very High Resoluuiion Imaging	AFRL/DEB _____	3- 6
DR Vincent P Giannamore	Xavier University of Louisiana , New Orleans , LA Environmentally-Benign synthesis of 1,5-Hexadiyne and Related Studies	AFRL/DEB _____	3- 7
DR Gurnam S Gill	Naval Postgraduate School , Monterey , CA Partitioning of Power Aperature Product of Space Based Radar	AFRL/VSS _____	3- 8
DR Robert J Hinde	Univ of Tennessee , Knoxville , TN Computational Aspects of the Spectral Theory of Physical and Chemical Binding	AFRL/DEB _____	3- 9
DR Martin A Hunter	Holy Cross College , Worcester , MA Reaction of Electronically-Excited Nitrogen Atoms with Molecular Oxygen	AFRL/VSF _____	3- 10
DR Brian D Jeffs	Brigham Young University , Provo , UT Deterministic Methods for Blind Restoration of Adaptive Optics Images of Space Objects	AFRL/DES _____	3- 11

SRP Final Report Table of Contents

Author	University/Institution Report Title	Phillips Laboratory Directorate	Vol-Page
DR Donald J Leo	Virginia Tech , Blacksburg , VA self-Sensing Techniqr for Active Acoustic Attenuation	AFRL/VSD _____	3- 12
DR M. Arfin K Lodhi	Texas Tech University , Lubbock , TX Effect of Materials and Design Variations on Amtec Cell Losses	AFRL/VSD _____	3- 13
DR John P McHugh	University of New Hampshire , Durham , NH A Splitting Technique for the anelastic equations in atmospheric physics.	AFRL/VSB _____	3- 14
DR Stanly L Steinberg	University of New Mexico , Albuquerque , NM Lie-Algebraic Representations of Product Intgrals of Variable Matrices	AFRL/DEH _____	3- 15

SRP Final Report Table of Contents

Author	University/Institution Report Title	Rome Laboratory Directorate	Vol-Page
DR Ercument Arvas	Syracuse University, Syracuse, NY Design of a Microwave-To-Optical Link Amplifier For Radar Applications	AFRL/SDN	4- 1
DR Milica Barjaktarovic	Wilkes University, Wilkes Barre, PA Information Protection Tools and Methods	AFRL/IFG	4- 2
DR Stella N Batalama	SUNY Buffalo, Buffalo, NY Outlier Resistant DS-SS Signal Processing	AFRL/IFG	4- 3
DR Digendra K Das	SUNYIT, Utica, NY Modeling and Simulation of Memos Resonators	AFRL/IFT	4- 4
DR Venugopala R Dasigi	, Marietta, GA Toward an Architecture For A Global Information Base	AFRL/CA-I	4- 5
DR Kaliappan Gopalan	Purdue Research Foundation, West Lafayette, IN Amplitude and Frequency Modulation Characteristics of Stressed Speech	AFRL/IFE	4- 6
DR Donald L Hung	Washington State University, Richland, WA A Study on Accelerating the Ray/Triangular-Facet Intersection Computation in Xpatch	AFRL/IFSA	4- 7
DR Adam Lutoborski	Syracuse University, Syracuse, NY On a wavelet-based method of watermarking digital images	AFRL/IFE	4- 8
DR Brajendra N Panda	University of North Dakota, Grand Forks, ND A Model to Analyze Sensor Data For Detection of Multi-Source Attacks	AFRL/IFG	4- 9
DR Jerry L Potter	Kent State University, Kent, OH Architectures for Knowledge Bases	AFRL/IFT	4- 10
DR Salahuddin Qazi	NY Coll of Tech Utica/Rome, Utica, NY Modeling and Implementation of Low Data Rate Modem Using Matlab	AFRL/IFG	4- 11

SRP Final Report Table of Contents

Author	University/Institution Report Title	Rome Laboratory Directorate	Vol-Page
DR Richard R Schultz	University of North Dakota , Grand Forks , ND Image Registration Algorithm Based on the Projective Transformation Model	AFRL/IFE _____	4- 12
DR Kalpathi R Subramanian	University of N. C.- Charlotte , Charlotte , NC Enhancements to Cubeworld	AFRL/IFSA _____	4- 13
DR Shambhu J Upadhyaya	SUNY Buffalo , Buffalo , NY a Distributed Concurrent Intrusion Detection Scheme Based on Assertions	AFRL/IFG _____	4- 14
DR Robert E Yantorno	Temple University , Philadelphia , PA Co-Channel Speech and Speaker Identification Study	AFRL/IFE _____	4- 15

SRP Final Report Table of Contents

Author	University/Institution Report Title	Wright Laboratory Directorate	Vol-Page
DR Farid Ahmed	Penn State Uni-Erie , Erie , PA Multiresolutional Information Feature for Dynamic Change Detecton in image Sequences	AFRL/SNA _____	5- 1
DR Kevin D Belfield	University of Central Florida , Orlando , FL Synthesis of 7-Benzothiazol-2YL-9,9-Didecylfluorene-2-Ylamine a versatile Intermediate for a New Ser	AFRL/ML _____	5- 2
DR Daniel D Bombick	Wright State University , Dayton , OH	AFRL/PRS _____	5- 3
DR Frank M Brown	University of Kansas , Lawrence , KS Recognizing Linearities In Manterials Databases	AFRL/ML _____	5- 4
DR Gregory A Buck	S Dakota School of Mines/Tech , Rapid City , SD Characterization of Acoustic Sources for Hypersonic Receptivity Research	AFRL/VAA _____	5- 5
DR Joe G Chow	Florida International Univ , Miami , FL Some Critical Issues of The Next Generation Transparency Program	AFRL/VAV _____	5- 6
DR Peter J Disimile	University of Cincinnati , Cincinnati , OH Documentation of the Airflow Patterns within and aircraft Engine Nacelle Simulator	AFRL/VAV _____	5- 7
DR Numan S Dogan	Tuskegee University , Tuskegee , AL Sensors for Focal Plane Array Passive Millimeter-Wave Imaging	AFRL/MN _____	5- 8
DR James M Fragomeni	Ohio University , Athens , OH Mechanical Strength Modeling of Particle strengthened Nickel-Aluminum Alloys Strengthened By Interme	AFRL/ML _____	5- 9
DR Zewdu Gebeyehu	Tuskegee University , Tuskegee , AL Synthesis & Characterization of Metal-Thioacid & Dihydrogen Phosphate Complexes Useful as Nonlinear	AFRL/MLP _____	5- 10
DR Patrick C Gilcrease	University of Wyoming , Laramie , WY Biocatalysis of Biphenyl and Diphenylacetylene to Synthesize Polymer Precursors	AFRL/ML _____	5- 11

SRP Final Report Table of Contents

Author	University/Institution Report Title	Wright Laboratory Directorate	Vol-Page
DR David E Hudak	Ohio Northern University , Ada , OH Permanence Modeling and Scalability Analysis of the Navier-Stokes Solver FDL3DI Across Multiple Platfo	AFRL/VAA _____	5- 12
DR William P Johnson	University of Utah , Salt Lake City , UT Sorption of a Non-Ionic Suractant Versus a Dissolved Humic Substance to a Low Orgnaic Carbon Soil	AFRL/ML _____	5- 13
DR Jeffrey D Johnson	University of Toledo , Toledo , OH Using Neural Networks to Control a Tailless Fighter Aircraft	AFRL/VAC _____	5- 14
DR Jayanta S Kapat	University of Central Florida , Orlando , FL Fuel-Air Heat Exhcngner For Cooled Cooling Air Systems with Fuel-Mist and Air-Jet Impingement	AFRL/PRT _____	5- 15
DR Vikram Kapila	Polytechnic Inst of New York , Brooklyn , NY Spacecraft Formation Flying: A Survey	AFRL/VAC _____	5- 16
DR Kenneth D Kihm	Texas Engineering Experiment Station , College Station , TX Micro-Scale Visualization of Thin Meniscus & Capillary Pore Flows of Capillary-Driven Heat Transfer	AFRL/VAV _____	5- 17
DR Lok C Lew Yan Voon	Worcester Polytechnic Inst , Worcester , MA Many-Body Theory of Quantum-Well Gain Spectra	AFRL/SND _____	5- 18
DR Rongxing Li	Ohio State University , Columbus , OH A Study fo Referencing Issues in Multiplatform and multisensor Based Object Location	AFRL/SNA _____	5- 19
DR Chun-Shin Lin	Univ of Missouri - Columbia , Columbia , MO Sensor Fusion w/Passive Millimeter Wave & Laser Radar for Target Detection	AFRL/MN _____	5- 20
DR Chaoqun Liu	Louisiana Tech University , Ruston , LA Boundary Conditions for Direct Numerical Simulation of Turbulent Flow	AFRL/VAA _____	5- 21
DR Carl E Mungan	University of Florida , Pensacola , FL Bidirectional Reflectance Distr. Functions Describing Firts-Surface Scattering	AFRL/MN _____	5- 22

SRP Final Report Table of Contents

Author	University/Institution Report Title	Wright Laboratory Directorate	Vol-Page
DR Amod A Ogale	Clemson University , Clemson , SC Characterization of Microstructure Evolution in Pitch-Based Carbon Fibers During Heat Treatment	AFRL/ML	5- 23
DR Carlos R Ortiz	Universidad Politecnica de Puerto Rico , Hato Rey , PR Simulation of the Antenna Pattern of Arbitrarily Oriented Very Large Phase/Time-Delay Scanned Antenn	AFRL/SNR	5- 24
DR Ramana M Pidaparti	Indiana U-Purdue at Indianap , Indianapolis , IN Flutter Prediction Methods for Aeroelastic Design Optimization	AFRL/VAS	5- 25
DR Stephen E Saddow	Mississippi State University , Mississippi State , MS Characterization of BN-Doped SiC Epitaxial Layers	AFRL/PRP	5- 26
DR Rathinam P Selvam	Univ of Arkansas , Fayetteville , AR Computer Modelling of Nonlinear Viscous Panel Flutter	AFRL/VAA	5- 27
DR Paavo Sepri	Florida Inst of Technology , Melbourne , FL A computational Study of Turbine Blade Interactions with Cylinder Wakes at Various Reynolds Numbers	AFRL/PRT	5- 28
DR Mo-How H Shen	Ohio State University , Columbus , OH Development of a Probabilistic Assessment Framework for High Cycle Fatigue Failures of gas Turbine E	AFRL/ML	5- 29
DR Hongchi Shi	Univ of Missouri - Columbia , Columbia , MO A Study of Models and Tools for Programming the VGI Parallel Computer	AFRL/MN	5- 30
DR Donald J Silversmith	Wayne State University , Detroit , MI Joule Heating Simulation of Poly-Silicon Thermal Micro-Actuators	AFRL/SNH	5- 31
DR Mehrdad Soumekh	SUNY Buffalo , Amherst , NY Alias-Free Processing of P-3 SAR Data	AFRL/SNR	5- 32
DR Joseph W Tedesco	Auburn University , Auburn , AL HIGH Velocity Penetration of Layered Grout Targets	AFRL/MN	5- 33

SRP Final Report Table of Contents

Author	University/Institution Report Title	Wright Laboratory Directorate	Vol-Page
DR Mitch J Wolff	Wright State University , Dayton , OH Enhancements to A Driect Aeroclastic Stability Computational Model	AFRL/VAS _____	5- 34
DR Jeffrey L Young	University of Idaho , Moscow , ID A Detrailed Study of the Numerical Properties of FDTD Algorithms for Dispersive Media	AFRL/VAA _____	5- 35

SRP Final Report Table of Contents

<u>Author</u>	<u>University/Institution</u> <u>Report Title</u>	<u>Laboratory</u> <u>Directorate</u>	<u>Vol-Page</u>
DR F. N. Albahadily	University of Central Oklahoma, Edmond, OH Effect of Environmental Variables on Aging Aircraft	OCALC _____	6 - 1
MS Shelia K Barnett	Mercer Univ, Macon, GA A Study of Scheduling and Tracking of Parts in the Plating Shop at Warner Robins Air Logistics Center	WRALC/TI _____	6 - 2
DR Ryan R Dupont	Utah State University, Logan, UT Natural Attenuation Evaluation Summary for a Chlorinated Solvent Plume, OUL, Hill AFB, Utah	OOALC/E _____	6 - 3
DR Carl L Enloe	James Madison Univ, Harrisonburg, VA A Device for Experimental Measurements of Elelctrostatic Shielding in a Spatially Non-Uniform Plasma	HQUSAF/D _____	6 - 4
DR Mark R Fisher	Southern Polytechnic State University, Marietta, GA Neural Network Control of Wind Tunnels for Cycle Time Reduction	AEDC _____	6 - 5
DR Sheng-Jen Hsieh	Pan American University, Edinbrg, TX Thermal Signature for Circuit Card Fault Identification	SAALC/TI _____	6 - 6
DR Suk B Kong	Incarnate Word College, San Antonio, TX Studies on The Amphetamine Derivatives and Analytical Standards	WHMC/59 _____	6 - 7
DR Kevin M Lyons	North Carolina State U-Raleigh, Raleigh, NC Filtered-Rayleigh Scattering in Reacting and Non-Reacting Flow	AEDC _____	6 - 8

1. INTRODUCTION

The Summer Research Program (SRP), sponsored by the Air Force Office of Scientific Research (AFOSR), offers paid opportunities for university faculty, graduate students, and high school students to conduct research in U.S. Air Force research laboratories nationwide during the summer.

Introduced by AFOSR in 1978, this innovative program is based on the concept of teaming academic researchers with Air Force scientists in the same disciplines using laboratory facilities and equipment not often available at associates' institutions.

The Summer Faculty Research Program (SFRP) is open annually to approximately 150 faculty members with at least two years of teaching and/or research experience in accredited U.S. colleges, universities, or technical institutions. SFRP associates must be either U.S. citizens or permanent residents.

The Graduate Student Research Program (GSRP) is open annually to approximately 100 graduate students holding a bachelor's or a master's degree; GSRP associates must be U.S. citizens enrolled full time at an accredited institution.

The High School Apprentice Program (HSAP) annually selects about 125 high school students located within a twenty mile commuting distance of participating Air Force laboratories.

AFOSR also offers its research associates an opportunity, under the Summer Research Extension Program (SREP), to continue their AFOSR-sponsored research at their home institutions through the award of research grants. In 1994 the maximum amount of each grant was increased from \$20,000 to \$25,000, and the number of AFOSR-sponsored grants decreased from 75 to 60. A separate annual report is compiled on the SREP.

The numbers of projected summer research participants in each of the three categories and SREP "grants" are usually increased through direct sponsorship by participating laboratories.

AFOSR's SRP has well served its objectives of building critical links between Air Force research laboratories and the academic community, opening avenues of communications and forging new research relationships between Air Force and academic technical experts in areas of national interest, and strengthening the nation's efforts to sustain careers in science and engineering. The success of the SRP can be gauged from its growth from inception (see Table 1) and from the favorable responses the 1997 participants expressed in end-of-tour SRP evaluations (Appendix B).

AFOSR contracts for administration of the SRP by civilian contractors. The contract was first awarded to Research & Development Laboratories (RDL) in September 1990. After completion of the 1990 contract, RDL (in 1993) won the recompetition for the basic year and four 1-year options.

2. PARTICIPATION IN THE SUMMER RESEARCH PROGRAM

The SRP began with faculty associates in 1979; graduate students were added in 1982 and high school students in 1986. The following table shows the number of associates in the program each year.

YEAR	SRP Participation, by Year			TOTAL
	SFRP	GSRP	HSAP	
1979	70			70
1980	87			87
1981	87			87
1982	91	17		108
1983	101	53		154
1984	152	84		236
1985	154	92		246
1986	158	100	42	300
1987	159	101	73	333
1988	153	107	101	361
1989	168	102	103	373
1990	165	121	132	418
1991	170	142	132	444
1992	185	121	159	464
1993	187	117	136	440
1994	192	117	133	442
1995	190	115	137	442
1996	188	109	138	435
1997	148	98	140	427
1998	85	40	88	213

Beginning in 1993, due to budget cuts, some of the laboratories weren't able to afford to fund as many associates as in previous years. Since then, the number of funded positions has remained fairly constant at a slightly lower level.

3. RECRUITING AND SELECTION

The SRP is conducted on a nationally advertised and competitive-selection basis. The advertising for faculty and graduate students consisted primarily of the mailing of 8,000 52-page SRP brochures to chairpersons of departments relevant to AFOSR research and to administrators of grants in accredited universities, colleges, and technical institutions. Historically Black Colleges and Universities (HBCUs) and Minority Institutions (MIs) were included. Brochures also went to all participating USAF laboratories, the previous year's participants, and numerous individual requesters (over 1000 annually).

RDL placed advertisements in the following publications: *Black Issues in Higher Education*, *Winds of Change*, and *IEEE Spectrum*. Because no participants list either *Physics Today* or *Chemical & Engineering News* as being their source of learning about the program for the past several years, advertisements in these magazines were dropped, and the funds were used to cover increases in brochure printing costs.

High school applicants can participate only in laboratories located no more than 20 miles from their residence. Tailored brochures on the HSAP were sent to the head counselors of 180 high schools in the vicinity of participating laboratories, with instructions for publicizing the program in their schools.

High school students selected to serve at Wright Laboratory's Armament Directorate (Eglin Air Force Base, Florida) serve eleven weeks as opposed to the eight weeks normally worked by high school students at all other participating laboratories.

Each SFRP or GSRP applicant is given a first, second, and third choice of laboratory. High school students who have more than one laboratory or directorate near their homes are also given first, second, and third choices.

Laboratories make their selections and prioritize their nominees. AFOSR then determines the number to be funded at each laboratory and approves laboratories' selections.

Subsequently, laboratories use their own funds to sponsor additional candidates. Some selectees do not accept the appointment, so alternate candidates are chosen. This multi-step selection procedure results in some candidates being notified of their acceptance after scheduled deadlines. The total applicants and participants for 1998 are shown in this table.

1998 Applicants and Participants			
PARTICIPANT CATEGORY	TOTAL APPLICANTS	SELECTEES	DECLINING SELECTEES
SFRP	382	85	13
(HBCU/MI)	(0)	(0)	(0)
GSRP	130	40	7
(HBCU/MI)	(0)	(0)	(0)
HSAP	328	88	22
TOTAL	840	213	42

4. SITE VISITS

During June and July of 1998, representatives of both AFOSR/NI and RDL visited each participating laboratory to provide briefings, answer questions, and resolve problems for both laboratory personnel and participants. The objective was to ensure that the SRP would be as constructive as possible for all participants. Both SRP participants and RDL representatives found these visits beneficial. At many of the laboratories, this was the only opportunity for all participants to meet at one time to share their experiences and exchange ideas.

5. HISTORICALLY BLACK COLLEGES AND UNIVERSITIES AND MINORITY INSTITUTIONS (HBCU/MIs)

Before 1993, an RDL program representative visited from seven to ten different HBCU/MIs annually to promote interest in the SRP among the faculty and graduate students. These efforts were marginally effective, yielding a doubling of HBCU/MI applicants. In an effort to achieve AFOSR's goal of 10% of all applicants and selectees being HBCU/MI qualified, the RDL team decided to try other avenues of approach to increase the number of qualified applicants. Through the combined efforts of the AFOSR Program Office at Bolling AFB and RDL, two very active minority groups were found, HACU (Hispanic American Colleges and Universities) and AISES (American Indian Science and Engineering Society). RDL is in communication with representatives of each of these organizations on a monthly basis to keep up with their activities and special events. Both organizations have widely-distributed magazines/quarterlies in which RDL placed ads.

Since 1994 the number of both SFRP and GSRP HBCU/MI applicants and participants has increased ten-fold, from about two dozen SFRP applicants and a half dozen selectees to over 100 applicants and two dozen selectees, and a half-dozen GSRP applicants and two or three selectees to 18 applicants and 7 or 8 selectees. Since 1993, the SFRP had a two-fold applicant increase and a two-fold selectee increase. Since 1993, the GSRP had a three-fold applicant increase and a three to four-fold increase in selectees.

In addition to RDL's special recruiting efforts, AFOSR attempts each year to obtain additional funding or use leftover funding from cancellations the past year to fund HBCU/MI associates.

SRP HBCU/MI Participation, By Year				
YEAR	SFRP		GSRP	
	Applicants	Participants	Applicants	Participants
1985	76	23	15	11
1986	70	18	20	10
1987	82	32	32	10
1988	53	17	23	14
1989	39	15	13	4
1990	43	14	17	3
1991	42	13	8	5
1992	70	13	9	5
1993	60	13	6	2
1994	90	16	11	6
1995	90	21	20	8
1996	119	27	18	7

6. SRP FUNDING SOURCES

Funding sources for the 1998 SRP were the AFOSR-provided slots for the basic contract and laboratory funds. Funding sources by category for the 1998 SRP selected participants are shown here.

1998 SRP FUNDING CATEGORY	SFRP	GSRP	HSAP
AFOSR Basic Allocation Funds	67	38	75
USAF Laboratory Funds	17	2	13
Slots Added by AFOSR (Leftover Funds)	0	0	0
HBCU/MI By AFOSR (Using Procured Addn'l Funds)	0	0	N/A
TOTAL	84	40	88

7. COMPENSATION FOR PARTICIPANTS

Compensation for SRP participants, per five-day work week, is shown in this table.

1998 SRP Associate Compensation

PARTICIPANT CATEGORY	1991	1992	1993	1994	1995	1996	1997	1998
Faculty Members	\$690	\$718	\$740	\$740	\$740	\$770	\$770	\$793
Graduate Student (Master's Degree)	\$425	\$442	\$455	\$455	\$455	\$470	\$470	\$484
Graduate Student (Bachelor's Degree)	\$365	\$380	\$391	\$391	\$391	\$400	\$400	\$412
High School Student (First Year)	\$200	\$200	\$200	\$200	\$200	\$200	\$200	\$200
High School Student (Subsequent Years)	\$240	\$240	\$240	\$240	\$240	\$240	\$240	\$240

The program also offered associates whose homes were more than 50 miles from the laboratory an expense allowance (seven days per week) of \$52/day for faculty and \$41/day for graduate students. Transportation to the laboratory at the beginning of their tour and back to their home destinations at the end was also reimbursed for these participants. Of the combined SFRP and GSRP associates, 65 % claimed travel reimbursements at an average round-trip cost of \$730.

Faculty members were encouraged to visit their laboratories before their summer tour began. All costs of these orientation visits were reimbursed. Forty-three percent (85 out of 188) of faculty associates took orientation trips at an average cost of \$449. By contrast, in 1993, 58 % of SFRP associates elected to take an orientation visits at an average cost of \$685; that was the highest percentage of

associates opting to take an orientation trip since RDL has administered the SRP, and the highest average cost of an orientation trip.

Program participants submitted biweekly vouchers countersigned by their laboratory research focal point, and RDL issued paychecks so as to arrive in associates' hands two weeks later.

This is the third year of using direct deposit for the SFRP and GSRP associates. The process went much more smoothly with respect to obtaining required information from the associates, about 15 % of the associates' information needed clarification in order for direct deposit to properly function as opposed to 7 % from last year. The remaining associates received their stipend and expense payments via checks sent in the US mail.

HSAP program participants were considered actual RDL employees, and their respective state and federal income tax and Social Security were withheld from their paychecks. By the nature of their independent research, SFRP and GSRP program participants were considered to be consultants or independent contractors. As such, SFRP and GSRP associates were responsible for their own income taxes, Social Security, and insurance.

8. CONTENTS OF THE 1998 REPORT

The complete set of reports for the 1998 SRP includes this program management report (Volume 1) augmented by fifteen volumes of final research reports by the 1998 associates, as indicated below:

1998 SRP Final Report Volume Assignments

LABORATORY	SFRP	GSRP	HSAP
Armstrong	2	7	12
Phillips	3	8	13
Rome	4	9	14
Wright	5A, 5B	10	15
AEDC, ALCs, USAFA, WHMC	6	11	

APPENDIX A – PROGRAM STATISTICAL SUMMARY

A. Colleges/Universities Represented

Selected SFRP associates represented 169 different colleges, universities, and institutions, GSRP associates represented 95 different colleges, universities, and institutions.

B. States Represented

SFRP -Applicants came from 47 states plus Washington D.C. Selectees represent 44 states.

GSRP - Applicants came from 44 states. Selectees represent 32 states.

HSAP - Applicants came from thirteen states. Selectees represent nine states.

Total Number of Participants	
SFRP	85
GSRP	40
HSAP	88
TOTAL	213

Degrees Represented			
	SFRP	GSRP	TOTAL
Doctoral	83	0	83
Master's	1	3	4
Bachelor's	0	22	22
TOTAL	186	25	109

SFRP Academic Titles	
Assistant Professor	36
Associate Professor	34
Professor	15
Instructor	0
Chairman	0
Visiting Professor	0
Visiting Assoc. Prof.	0
Research Associate	0
TOTAL	85

Source of Learning About the SRP		
Category	Applicants	Selectees
Applied/participated in prior years	177	47
Colleague familiar with SRP	104	24
Brochure mailed to institution	101	21
Contact with Air Force laboratory	101	39
<i>IEEE Spectrum</i>	12	1
<i>BIIHE</i>	4	0
Other source	117	30
TOTAL	616	162

APPENDIX B – SRP EVALUATION RESPONSES

1. OVERVIEW

Evaluations were completed and returned to RDL by four groups at the completion of the SRP. The number of respondents in each group is shown below.

Table B-1. Total SRP Evaluations Received

Evaluation Group	Responses
SFRP & GSRPs	100
HSAPs	75
USAF Laboratory Focal Points	84
USAF Laboratory HSAP Mentors	6

All groups indicate unanimous enthusiasm for the SRP experience.

The summarized recommendations for program improvement from both associates and laboratory personnel are listed below:

- A. Better preparation on the labs' part prior to associates' arrival (i.e., office space, computer assets, clearly defined scope of work).
- B. Faculty Associates suggest higher stipends for SFRP associates.
- C. Both HSAP Air Force laboratory mentors and associates would like the summer tour extended from the current 8 weeks to either 10 or 11 weeks; the groups state it takes 4-6 weeks just to get high school students up-to-speed on what's going on at laboratory. (Note: this same argument was used to raise the faculty and graduate student participation time a few years ago.)

2. 1998 USAF LABORATORY FOCAL POINT (LFP) EVALUATION RESPONSES

The summarized results listed below are from the 84 LFP evaluations received.

1. LFP evaluations received and associate preferences:

Table B-2. Air Force LFP Evaluation Responses (By Type)

Lab	Evals Recv'd	How Many Associates Would You Prefer To Get ?								(% Response)			
		SFRP				GSRP (w/Univ Professor)				GSRP (w/o Univ Professor)			
		0	1	2	3+	0	1	2	3+	0	1	2	3+
AEDC	0	-	-	-	-	-	-	-	-	-	-	-	-
WHMC	0	-	-	-	-	-	-	-	-	-	-	-	-
AL	7	28	28	28	14	54	14	28	0	86	0	14	0
USAF	1	0	100	0	0	100	0	0	0	0	100	0	0
PL	25	40	40	16	4	88	12	0	0	84	12	4	0
RL	5	60	40	0	0	80	10	0	0	100	0	0	0
WL	46	30	43	20	6	78	17	4	0	93	4	2	0
Total	84	32%	50%	13%	5%	80%	11%	6%	0%	73%	23%	4%	0%

LFP Evaluation Summary. The summarized responses, by laboratory, are listed on the following page. LFPs were asked to rate the following questions on a scale from 1 (below average) to 5 (above average).

2. LFPs involved in SRP associate application evaluation process:
 - a. Time available for evaluation of applications:
 - b. Adequacy of applications for selection process:
3. Value of orientation trips:
4. Length of research tour:
5.
 - a. Benefits of associate's work to laboratory:
 - b. Benefits of associate's work to Air Force:
6.
 - a. Enhancement of research qualifications for LFP and staff:
 - b. Enhancement of research qualifications for SFRP associate:
 - c. Enhancement of research qualifications for GSRP associate:
7.
 - a. Enhancement of knowledge for LFP and staff:
 - b. Enhancement of knowledge for SFRP associate:
 - c. Enhancement of knowledge for GSRP associate:
8. Value of Air Force and university links:
9. Potential for future collaboration:
10.
 - a. Your working relationship with SFRP:
 - b. Your working relationship with GSRP:
11. Expenditure of your time worthwhile:

(Continued on next page)

12. Quality of program literature for associate:
13. a. Quality of RDL's communications with you:
 b. Quality of RDL's communications with associates:
14. Overall assessment of SRP:

Table B-3. Laboratory Focal Point Responses to above questions

	<i>AEDC</i>	<i>AL</i>	<i>USAFA</i>	<i>PL</i>	<i>RL</i>	<i>WHMC</i>	<i>WL</i>
<i># Evals Recv'd</i>	0	7	1	14	5	0	46
<i>Question #</i>							
2	-	86 %	0 %	88 %	80 %	-	85 %
2a	-	4.3	n/a	3.8	4.0	-	3.6
2b	-	4.0	n/a	3.9	4.5	-	4.1
3	-	4.5	n/a	4.3	4.3	-	3.7
4	-	4.1	4.0	4.1	4.2	-	3.9
5a	-	4.3	5.0	4.3	4.6	-	4.4
5b	-	4.5	n/a	4.2	4.6	-	4.3
6a	-	4.5	5.0	4.0	4.4	-	4.3
6b	-	4.3	n/a	4.1	5.0	-	4.4
6c	-	3.7	5.0	3.5	5.0	-	4.3
7a	-	4.7	5.0	4.0	4.4	-	4.3
7b	-	4.3	n/a	4.2	5.0	-	4.4
7c	-	4.0	5.0	3.9	5.0	-	4.3
8	-	4.6	4.0	4.5	4.6	-	4.3
9	-	4.9	5.0	4.4	4.8	-	4.2
10a	-	5.0	n/a	4.6	4.6	-	4.6
10b	-	4.7	5.0	3.9	5.0	-	4.4
11	-	4.6	5.0	4.4	4.8	-	4.4
12	-	4.0	4.0	4.0	4.2	-	3.8
13a	-	3.2	4.0	3.5	3.8	-	3.4
13b	-	3.4	4.0	3.6	4.5	-	3.6
14	-	4.4	5.0	4.4	4.8	-	4.4

3. 1998 SFRP & GSRP EVALUATION RESPONSES

The summarized results listed below are from the 120 SFRP/GSRP evaluations received.

Associates were asked to rate the following questions on a scale from 1 (below average) to 5 (above average) - by Air Force base results and over-all results of the 1998 evaluations are listed after the questions.

1. The match between the laboratories research and your field:
2. Your working relationship with your LFP:
3. Enhancement of your academic qualifications:
4. Enhancement of your research qualifications:
5. Lab readiness for you: LFP, task, plan:
6. Lab readiness for you: equipment, supplies, facilities:
7. Lab resources:
8. Lab research and administrative support:
9. Adequacy of brochure and associate handbook:
10. RDL communications with you:
11. Overall payment procedures:
12. Overall assessment of the SRP:
13.
 - a. Would you apply again?
 - b. Will you continue this or related research?
14. Was length of your tour satisfactory?
15. Percentage of associates who experienced difficulties in finding housing:
16. Where did you stay during your SRP tour?
 - a. At Home:
 - b. With Friend:
 - c. On Local Economy:
 - d. Base Quarters:
17. Value of orientation visit:
 - a. Essential:
 - b. Convenient:
 - c. Not Worth Cost:
 - d. Not Used:

SFRP and GSRP associate's responses are listed in tabular format on the following page.

Table B-4. 1997 SFRP & GSRP Associate Responses to SRP Evaluation

	Arnold	Brooks	Edwards	Eglin	Griffis	Hanscom	Kelly	Kirtland	Lackland	Robins	Tyndall	WPAFB	average
# res	6	48	6	14	31	19	3	32	1	2	18	85	257
1	4.8	4.4	4.6	4.7	4.4	4.9	4.6	4.6	5.0	5.0	4.0	4.7	4.6
2	5.0	4.6	4.1	4.9	4.7	4.7	5.0	4.7	5.0	5.0	4.6	4.8	4.7
3	4.5	4.4	4.0	4.6	4.3	4.2	4.3	4.4	5.0	5.0	4.5	4.3	4.4
4	4.3	4.5	3.8	4.6	4.4	4.4	4.3	4.6	5.0	4.0	4.4	4.5	4.5
5	4.5	4.3	3.3	4.8	4.4	4.5	4.3	4.2	5.0	5.0	3.9	4.4	4.4
6	4.3	4.3	3.7	4.7	4.4	4.5	4.0	3.8	5.0	5.0	3.8	4.2	4.2
7	4.5	4.4	4.2	4.8	4.5	4.3	4.3	4.1	5.0	5.0	4.3	4.3	4.4
8	4.5	4.6	3.0	4.9	4.4	4.3	4.3	4.5	5.0	5.0	4.7	4.5	4.5
9	4.7	4.5	4.7	4.5	4.3	4.5	4.7	4.3	5.0	5.0	4.1	4.5	4.5
10	4.2	4.4	4.7	4.4	4.1	4.1	4.0	4.2	5.0	4.5	3.6	4.4	4.3
11	3.8	4.1	4.5	4.0	3.9	4.1	4.0	4.0	3.0	4.0	3.7	4.0	4.0
12	5.7	4.7	4.3	4.9	4.5	4.9	4.7	4.6	5.0	4.5	4.6	4.5	4.6
Numbers below are percentages													
13a	83	90	83	93	87	75	100	81	100	100	100	86	87
13b	100	89	83	100	94	98	100	94	100	100	100	94	93
14	83	96	100	90	87	80	100	92	100	100	70	84	88
15	17	6	0	33	20	76	33	25	0	100	20	8	39
16a	-	26	17	9	38	23	33	4	-	-	-	30	
16b	100	33	-	40	-	8	-	-	-	-	36	2	
16c	-	41	83	40	62	69	67	96	100	100	64	68	
16d	-	-	-	-	-	-	-	-	-	-	-	0	
17a	-	33	100	17	50	14	67	39	-	50	40	31	35
17b	-	21	-	17	10	14	-	24	-	50	20	16	16
17c	-	-	-	-	10	7	-	-	-	-	-	2	3
17d	100	46	-	66	30	69	33	37	100	-	40	51	46

4. 1998 USAF LABORATORY HSAP MENTOR EVALUATION RESPONSES

Not enough evaluations received (5 total) from Mentors to do useful summary.

5. 1998 HSAP EVALUATION RESPONSES

The summarized results listed below are from the 23 HSAP evaluations received.

HSAP apprentices were asked to rate the following questions on a scale from
1 (below average) to 5 (above average)

1. Your influence on selection of topic/type of work.
2. Working relationship with mentor, other lab scientists.
3. Enhancement of your academic qualifications.
4. Technically challenging work.
5. Lab readiness for you: mentor, task, work plan, equipment.
6. Influence on your career.
7. Increased interest in math/science.
8. Lab research & administrative support.
9. Adequacy of RDL's Apprentice Handbook and administrative materials.
10. Responsiveness of RDL communications.
11. Overall payment procedures.
12. Overall assessment of SRP value to you.
13. Would you apply again next year? Yes (92 %)
14. Will you pursue future studies related to this research? Yes (68 %)
15. Was Tour length satisfactory? Yes (82 %)

	Arnold	Brooks	Edwards	Eglin	Griffiss	Hanscom	Kirtland	Tyndall	WPAFB	Totals
# resp	5	19	7	15	13	2	7	5	40	113
1	2.8	3.3	3.4	3.5	3.4	4.0	3.2	3.6	3.6	3.4
2	4.4	4.6	4.5	4.8	4.6	4.0	4.4	4.0	4.6	4.6
3	4.0	4.2	4.1	4.3	4.5	5.0	4.3	4.6	4.4	4.4
4	3.6	3.9	4.0	4.5	4.2	5.0	4.6	3.8	4.3	4.2
5	4.4	4.1	3.7	4.5	4.1	3.0	3.9	3.6	3.9	4.0
6	3.2	3.6	3.6	4.1	3.8	5.0	3.3	3.8	3.6	3.7
7	2.8	4.1	4.0	3.9	3.9	5.0	3.6	4.0	4.0	3.9
8	3.8	4.1	4.0	4.3	4.0	4.0	4.3	3.8	4.3	4.2
9	4.4	3.6	4.1	4.1	3.5	4.0	3.9	4.0	3.7	3.8
10	4.0	3.8	4.1	3.7	4.1	4.0	3.9	2.4	3.8	3.8
11	4.2	4.2	3.7	3.9	3.8	3.0	3.7	2.6	3.7	3.8
12	4.0	4.5	4.9	4.6	4.6	5.0	4.6	4.2	4.3	4.5
Numbers below are percentages										
13	60%	95%	100%	100%	85%	100%	100%	100%	90%	92%
14	20%	80%	71%	80%	54%	100%	71%	80%	65%	68%
15	100%	70%	71%	100%	100%	50%	86%	60%	80%	82%

REF ID: A60001

**TEMPORAL CHARACTERISATION OF A SYNCHRONOUSLY-PUMPED PERIODICALLY-POLED
LITHIUM NIOBATE OPTICAL PARAMETRIC OSCILLATOR**

**Dr. Graham R. Allan
Associate Professor of Physics
Principal Scientist**

**New Mexico Highlands University
National Avenue
Las Vegas NM 87701**

now at

**Goddard Space Flight Center
NASA
Greenbelt Md. 20771**

**Final Report for:
Summer Faculty Research Program
Phillips Laboratory**

**Sponsored By:
Air Force Office of Scientific Research
Bolling Air Force Base
Washington DC**

&

Phillips Laboratory

August 98

TEMPORAL CHARACTERISATION OF A SYNCHRONOUSLY-PUMPED PERIODICALLY-POLED LITHIUM NIOBATE OPTICAL PARAMETRIC OSCILLATOR

[†]Graham R. Allan, Ph.D.
Associate Professor of Physics

New Mexico Highlands University
National Avenue
Las Vegas NM 87701

Abstract

The temporal characterisation of a synchronously-pumped periodically-poled Lithium Niobate optical parametric oscillator through OPO switch-on to steady state oscillation is presented. The micron-spatial and picosecond-temporal resolution of the detection system is used to explore the dynamics of the OPO through switch-on and build-up to steady state oscillation. The dynamical nature of the system, it is hypothesised, can be modeled by a simple damped harmonic oscillator. The oscillations are manifest in the pump transmission as energy flows between the pump field and the signal field. By varying the cavity losses through a variable reflectivity output coupler experimental evidence was collected to support this hypothesis. The characteristic times obtained for energy oscillation between the pump and signal field is found to be in the tens of Kilohertz with a decay constant of $\sim 20 \mu\text{s}$, indicating that the system can be empirically modeled using a slightly underdamped harmonic oscillator model.

[†] Dr. Allan is now with the
Lasers and Electro-Optics Branch
Goddard Space Flight Center
NASA
Greenbelt Md. 20771

TEMPORAL CHARACTERISATION OF A SYNCHRONOUSLY-PUMPED PERIODICALLY-POLED LITHIUM NIOBATE OPTICAL PARAMETRIC OSCILLATOR

Graham R. Allan, Ph.D.

Introduction

In this tour work concentrated on three main areas; experimental research into the dynamics of an OPO, finalising a paper for publication and training of personnel. The OPO work continues the temporal characterisation of a PPLN OPO and included the reconfiguration of the experimental system and familiarising new personnel with the data acquisition hardware and software. Work on the paper continues finalising diagrams and text.

An Optical Parametric Oscillator (OPO) is a device that uses a cavity to tune the response of a nonlinear optical medium to an intense laser input, the pump signal, so as to down-convert the frequency of the laser light into a lower-frequency output signal. OPO's are now well established as a means of generating widely-tunable coherent radiation in regions of the electro-magnetic spectrum which cannot be accessed directly via conventional coherent sources. However, a limitation of OPO's as a tool in nonlinear optics research is the low intensity of the output signal. A material which shows significant promise as a nonlinear optical medium in which to generate widely-tunable, coherent, near-infrared radiation efficiently with useful intensities is Periodically-Poled Lithium Niobate (PPLN). In addition to finding improved materials, deriving more power from an OPO also necessitates optimizing the design of the synchronously-pumped OPO cavity for a particular nonlinear material. Optimizing the OPO design requires detailed understanding of the dynamics involved in the forward conversion of the pump to signal. Since the process is a dynamical one, both temporal and spatial factor play a role.

The analysis of data presented in previous End of Tour Report¹ pertains to the steady-state region of the pump train only. This analysis indicates that the conversion efficiency is limited by back conversion of the signal to pump. This summer the work extended the measurements of the transient response during the period following switch-on but prior to steady-state equilibration. Preliminary analysis of this data reveals significant structure during turn-on. This new structure has a relatively long time scale, $> 50 \mu\text{s}$, whose presence signifies the importance of coupled cavity/nonlinear-medium dynamics in determining the system performance.

The research described herein is performed in the singly-resonant oscillator configuration, where the cavity Q is high for the signal frequency but low for the pump and idler. Feedback is achieved by circulating the signal within the OPO cavity and aligning for temporal and spatial overlap of the signal and pump within the PPLN crystal. This results in a gain in the signal strength for each pass

through the crystal. The gain mechanism is a $\chi^{(2)}$ nonlinear effect; maximum pump conversion takes place at the spatial and temporal center of the pump pulse (where the E-field is largest), which is the mode-locked output of a Nd:YAG laser operating at the fundamental frequency. To properly measure the OPO's dynamics the detection system must be fast compared to the temporal pulse-width and sample a small area compared to the beam diameter. Otherwise, one automatically measures an averaged output, integrated over space and time, and much of the information on the dynamics of the OPO may be lost.

Experimental

A schematic of the general experimental layout is shown in figure 1. A Coherent Antares ND:YAG operating at 1.064 μm in mode-locked quasi-cw mode produces an ~ 80 MHz train of pulses, nominally 100 ps FWHM at 20 W average power.

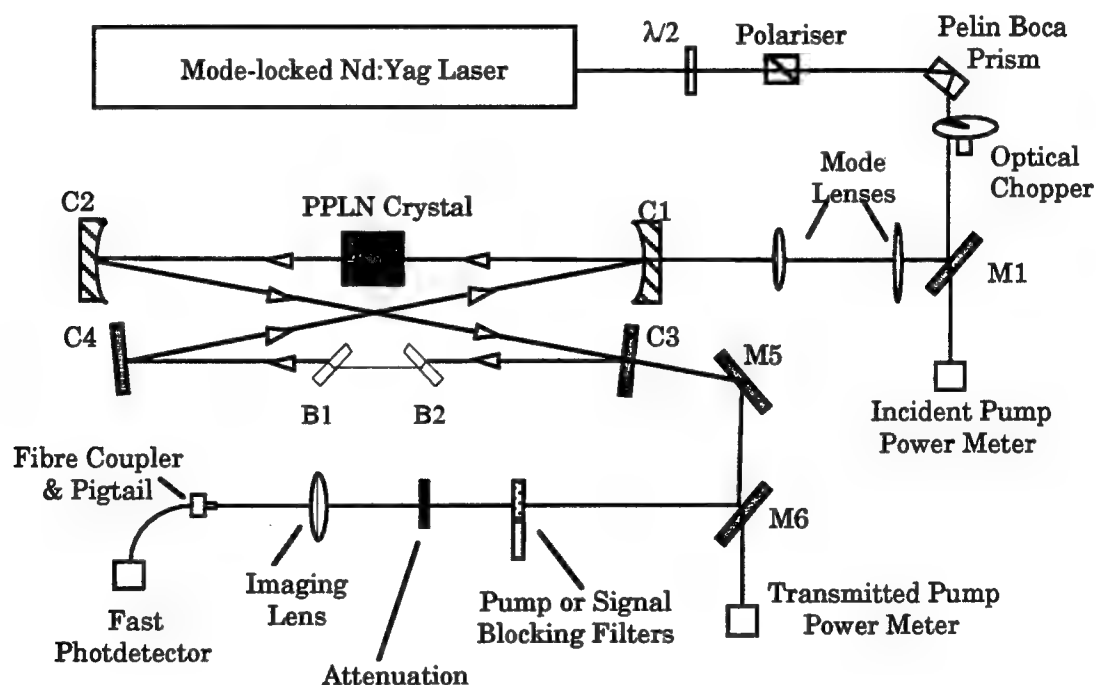


Figure 1. Schematic of the experimental layout showing the cavity geometry.

The pump power is injected into the OPO cavity through mirror C1, controlled by the halfwave plate and analyzer, and measured by a large area power meter at M1. The OPO cavity is a "bow-tie" ring resonator. The crystal is positioned centrally between mirrors C1 and C2, each of radius 75 cm, separated by 86 cm. The cavity is folded via flat reflectors C3 and C4. The optical path of the OPO is matched to the temporal-mode spacing of the pump using C4. The pump is spatially mode-matched to

the OPO cavity with an external telescope, producing a symmetrical Gaussian beam of $150\text{ }\mu\text{m}$ half-width ($1/e$).

The crystal is mounted in an externally-controlled oven to enable temperature tuning of the signal and idler frequencies. The crystal of LiNbO_3 is $15\text{ mm} \times 15\text{ mm} \times 500\text{ }\mu\text{m}$ thick, with a poling period of $29.75\text{ }\mu\text{m}$ through a central area of 13 mm diameter. The depleted pump, having passed through the crystal, is partially reflected by mirror C2 and leaves the resonator through C3. A large area power meter placed after a high reflectance $1.51\text{-}\mu\text{m}$ splitter, M6, monitors the transmitted pump power.

The cavity and crystal are optimized for operation at the signal wavelength of $1.51\text{ }\mu\text{m}$. The signal frequency is coupled from the cavity at mirrors C2 and C3. The fast temporally and spatially-resolving detection system follows in the diagnostic arm of the experiment. Depending on the measurement, a $1.51\text{ }\mu\text{m}$ or $1.064\text{ }\mu\text{m}$ transmission filter is inserted downstream of M6. The second harmonic of $1.51\text{ }\mu\text{m}$ is removed and attenuation, when needed, introduced.

A pair of intracavity glass plates, B1 and B2, mounted on rotation stages act as variable output couplers. The two plates were inserted with opposite tilt to cancel walkoff due to the refraction at the glass surfaces. Changing the angle of the plates alters the Fresnel losses and thus the outcoupling experienced by the circulating signal. The output coupling can be varied smoothly from as much as 18% output coupling, when the plates are near normal incidence, to less than 0.1% at Brewster's angle.

The pump is mechanically chopped at 40 Hz with a 100:1 duty cycle to minimize bulk heating of the crystal. This reduces the pump train to $250\text{ }\mu\text{s}$ bursts every 25 ms. The buildup to steady state in this resonator occurs in $\sim 50\text{ }\mu\text{s}$, and the chopper window is sufficiently wide to establish steady-state oscillation, where the single-pass conversion is balanced by round-trip cavity losses. By spatially and temporally resolving the transmitted pulse and signal during this $250\text{ }\mu\text{s}$ temporal window one can characterize the dynamics of the oscillator.

To this end a fast photodetector, a New Focus 10-410 with a rise time of 25 ps, is fiber coupled through a $25\text{ }\mu\text{m}$ diameter fiber at the experimental input and connected to an S-4 Tektronics sampling head. A narrow, variable, electronically-delayed sampling window of $\sim 20\text{ ps}$ width is set to record that temporal component of interest within the pulse. This output is then monitored as a function of temporal delay within the $250\text{ }\mu\text{s}$ chopper window. Spatial resolution is provided by a three-axis, computer-controlled positioning stage that has sub-micron accuracy and repeatability. This procedure enables temporal and spatial measurement of the transmitted pump through switch-on, steady-state operation and switch-off.

A dynamic background subtraction technique was employed to improve signal to noise while using the analog sampling head. This lessens significantly the susceptibility to baseline drift in the analog sampling scope. The sample head produces a DC voltage proportional to the detected signal within the bin. This signal is then sent to two boxcars, one is used to monitor the signal the other set to measure the back ground signal. The two boxcars time-constants and gains equated and the difference between the signals is free from background fluctuations.

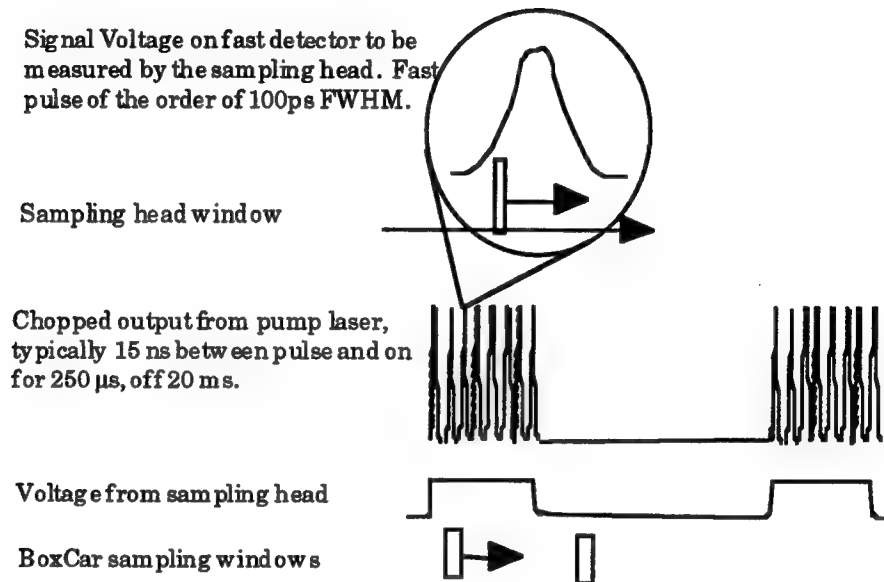


Figure 2. Dynamic Background Subtraction showing fast pulse, pulse train, output voltage from the sampling head and variable and fixed delay measuring windows.

On inspecting the temporal nature of the transmitted pump following turn-on, four distinct regions are discerned, as sketched qualitatively in figure 2: Full pump strength is attained when the beam clears the chopper window fully, within $\sim 1\mu\text{s}$ of encountering the edge; There then follows an increase in transmission of the pump until threshold, region B, where the pump depletion is at a minimum; Subsequently, in region C the transmitted pump undershoots the equilibrium value, and pump depletion is maximized; Finally, the transmitted signal rebounds slightly and approaches its equilibrium value asymptotically in the steady state region, D. There are several time constants inherent in the observed dynamics that characterize the system, and understanding their physical significance provides insight into the functioning of the OPO. The time constants are long compared to both the round-trip travel time in the cavity and the photon lifetime--the structure occurs over many thousands of round trips. The oscillations most likely represent energy transfer from pump to signal and back to pump, suggesting that a proper understanding of the transient dynamics of pump depletion is essential to optimizing OPO design.

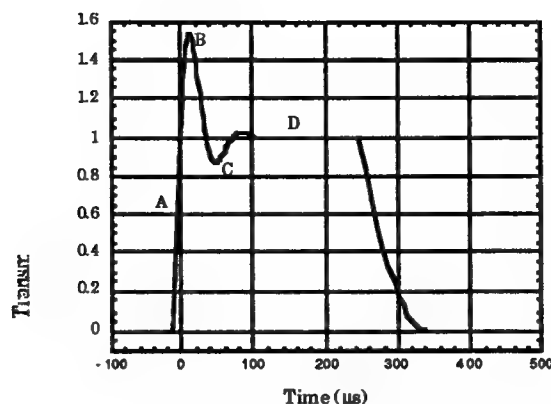


Figure 3. Qualitative sketch of the transmitted pump as a function of time in a synchronously pumped PPLN OPO. The chopper window is open for $\sim 250\mu\text{s}$ producing a temporally square quasi-cw input beam composed of the near Gaussian temporal and spatial mode locked pulses. Refer to text for labels.

The mechanism for pump depletion may be regarded as a process of energy conversion from the pump to the signal field, which is stored within the oscillator cavity. The temporal dependence of the transmitted pump is suggestive of a slightly under-damped resonant system that is ringing-down in response to a square-pulse stimulus. Tentatively, one can analyze the structure in terms of an LCR analogy: The transmitted intensity represents the system's response to the input, in this case a $250\mu\text{s}$, quasi-cw incident pulse wave train (square wave); The cavity corresponds to the LC component of the system, with "resistance" provided by cavity losses (including output coupling). Maintaining the resonant-circuit analogy, energy oscillates between the pump field and signal field just as it does between the magnetic field and electric field in the inductor and capacitor in the LCR circuit. Experimentally to measure these parameters and to extract time constants requires varying the output coupling and recording the temporal evolution of the transmitted pump through switchon to steady state.

Through analysis of the transmitted pump intensity as a function of output coupling one can extract several time constants that characterize the dynamics of the PPLN OPO system. Understanding the origins of these time constants will allow more efficient cavity design to optimize the output of the sync-pumped PPLN OPO's

Results and Discussion

To determine the characteristic time-constants of the OPO we monitored the transmitted pump power as a function of output coupling. The two intracavity glass plates change the cavity loss (output coupling) through the angular dependence of the Fresnel loss. As the plates' angles were adjusted the cavity length was tuned to maintain synchronism. Three output coupling losses were chosen arbitrarily, 2%, 4% and 16%. Figures 4a and 4b show a typical experimental plot of the

transmitted pump's temporal evolution through switch-on, steady state and switch-off, recorded at the spatial center of the transmitted pump pulse for 16% outcoupling. The fast axis (ps \times 10 timescale) is the time-resolved temporal profile of the transmitted pump. The slow axis (μ s \times 100 timescale) corresponds to time within the chopper window. The origin corresponds to the start of the fast pulse and the opening of the chopper window.

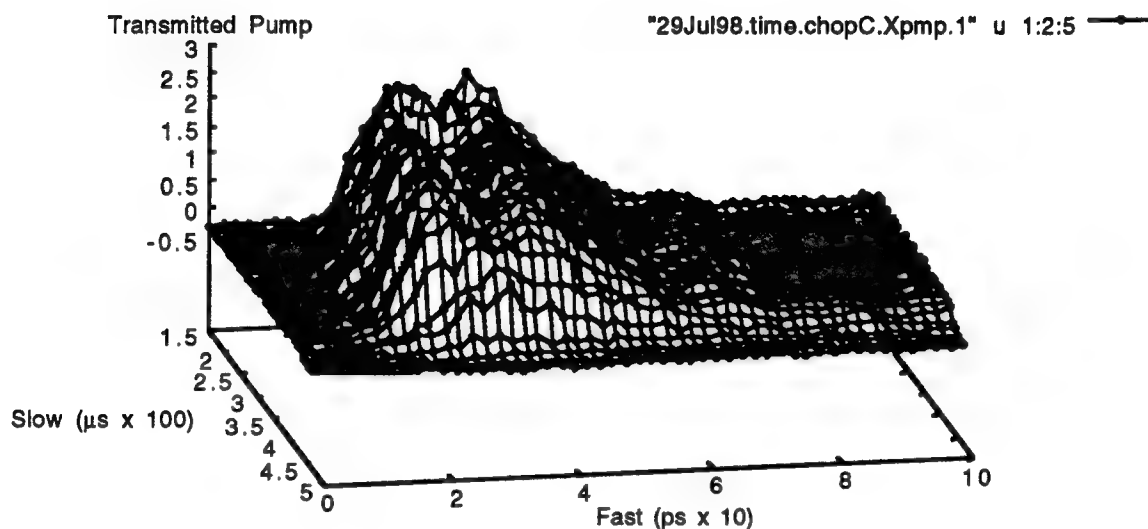


Figure 4a. Three dimensional plot of the transmitted pump as a function of time from switch-on to steady state operation to switch-off of the OPO. The center has been clearly depleted.

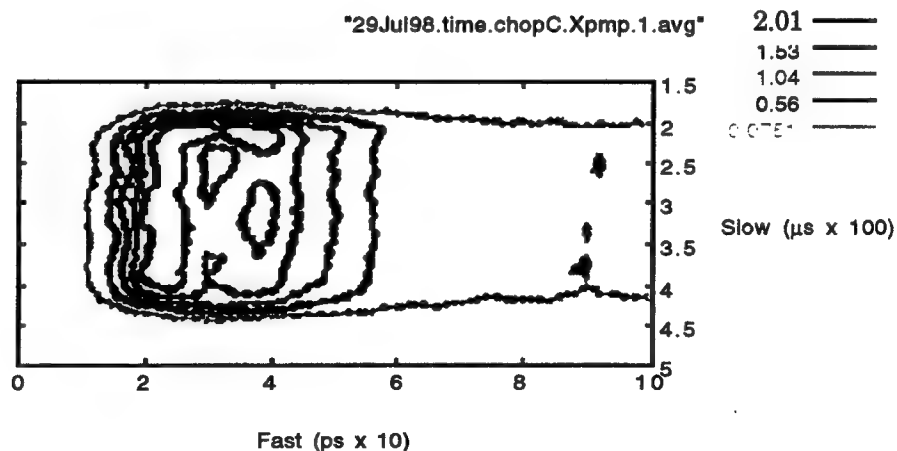


Figure 4b. Contour plot of the temporal evolution of the transmitted pump shown in Fig. 4a. The center portion has been clearly depleted.

The long tail in the fast axis with slight ringing is electronic noise generated within the fast detector. From this plot pump depletion can be clearly seen within the transmitted pump profile. Technical problems with the stability of the sampling head trigger leads to a slow drift in the delay as evident from figure 5a and 5b.

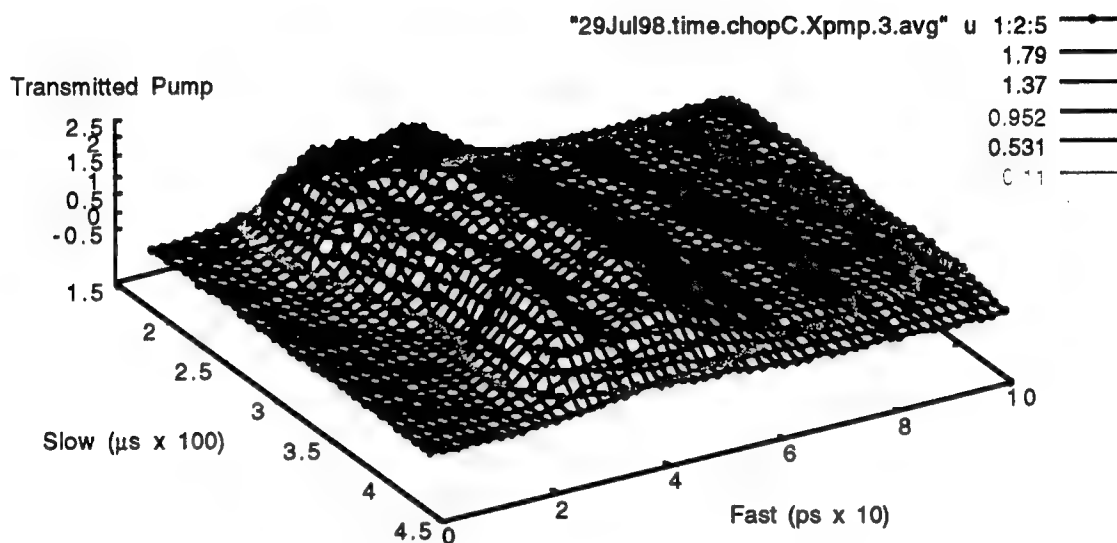


Figure 5a. Three dimensional plot of the temporal evolution of the transmitted pump as a function of time from swithcon showing drift in sample-head triggering. The pulse clearly shows pump depletion but with a drift in the start time of the pulse.

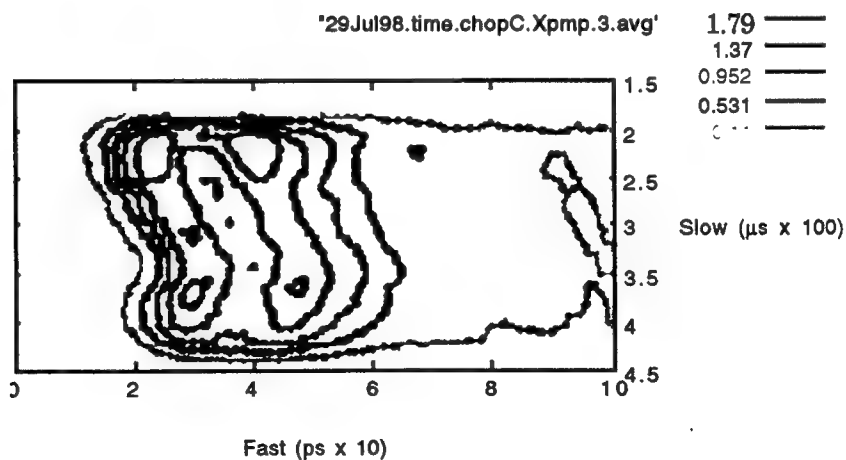


Figure 5b. Contour plot of the temporal evolution of the transmitted pump shown in Fig. 5a.

The slow variation in the turn-on of the fast profile is a result of drift. This data was "manually" re-centered and is shown below. This required writing software to split the data array and individually center the fast temporal profile while slightly clipping the data set to keep the dimensions of the arrays constant.

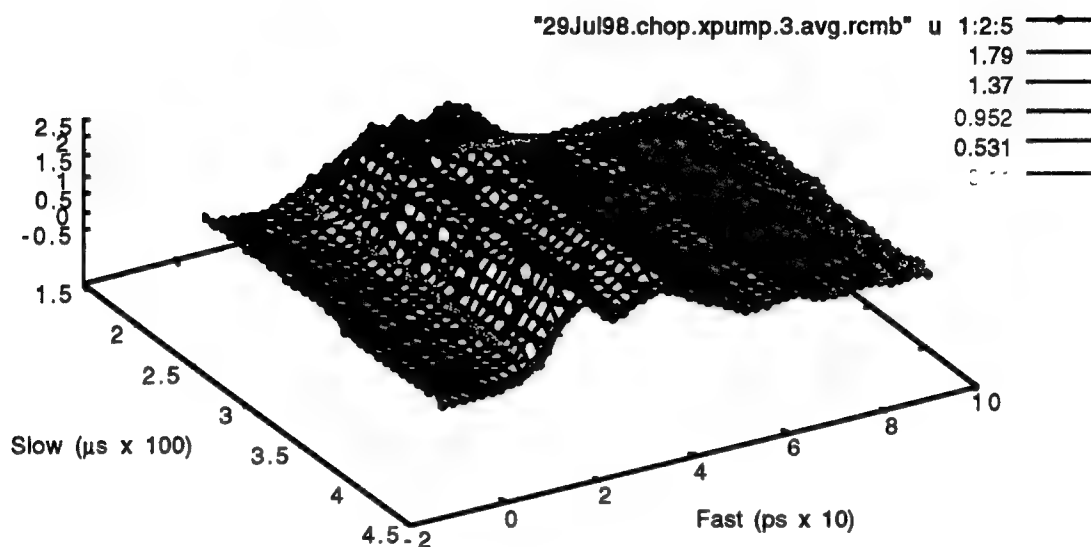


Figure 6a. Transmitted pump replotted after correcting for time base drift. The plot has been cut in the steady state regime and clearly shows pump depletion on the fast temporal profile.

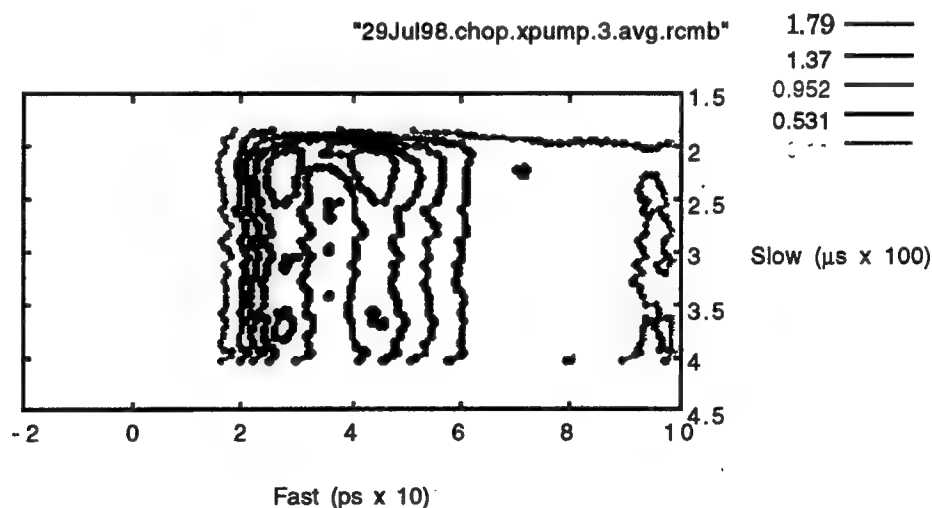


Figure 6b. Recentered fast profiles again showing pump depletion.

To determine the time constants the data had to be "sliced" along the slow temporal axis to produce temporal profiles through the chopper window. These plots then depict the temporal evolution of, the temporal center of the pump and reveal information on the dynamics of the OPO. Again the necessary code had to be written to take the data arrays and select elements at the same delay within the fast pulse but collected for all delays within the chopper window. Figure 7 plots the transmitted pump measured at the temporal and spatial center of the fast pulse as a function of time from switch on to steady state operation to the OPO for three different output coupling strengths, 2%, 4%, and 16%.

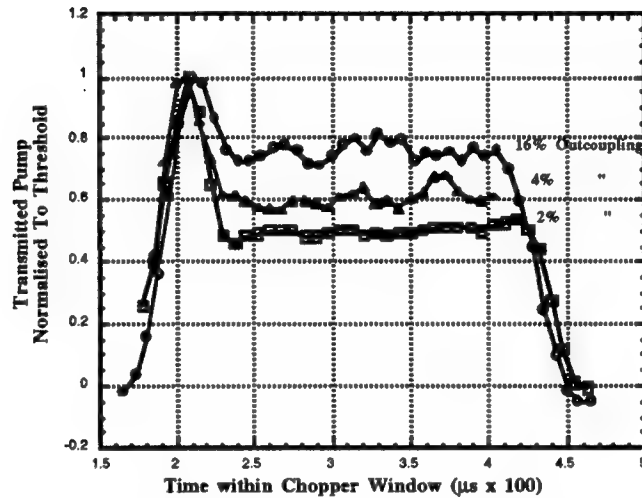


Figure 7. Transmitted pump as a function of time. The peak on the right corresponds to the transmitted pump before switch on of the OPO.

The data presented is normalised to the peak transmission at threshold. As can be seen from the figure the pump is quickly depleted and reaches a steady state value within $\sim 50\mu\text{s}$. The hypothesis is that the system behaves and can be modeled as a damped harmonic oscillator. A general solution of the differential equations describing a damped harmonic system is the product of a sine function and exponential decay. Figure 8 is a plot of the data with a fit to a slightly underdamped system.

$$(1 - A \cos(\omega t)) e^{(-t/\tau)}$$

where A is (1- steady state transmission), ω the period of oscillation, τ decay constant and t time. To reduce the number of fitting parameters all plots were normalised to the peak transmission at threshold. Figure 8 plots the data and fits where only the scaling factor "A" was change. From this fit it is evident that one period and one decay time does not fit all the data. The period fits the 16% data but the decay time tends to be too short to accurately fit the first minima but too long to fit the second minima. This may suggest two time constant but more likely the signal is lost in the noise.

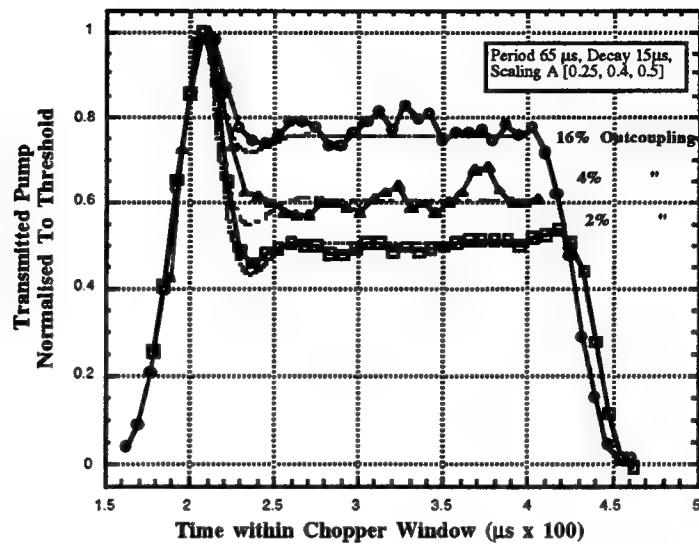


Figure 8. Transmitted pump as a function of time. The dashed curves are fits to the data.

Changing the scaling factor "A" only produces a less than adequate fit to the 4% and 2% coupling. Figure 9 plots fit where A, ω and τ were varied. Typical values reveal an oscillation between pump and signal in the tens of kilohertz with a decay time of 17 μ s.

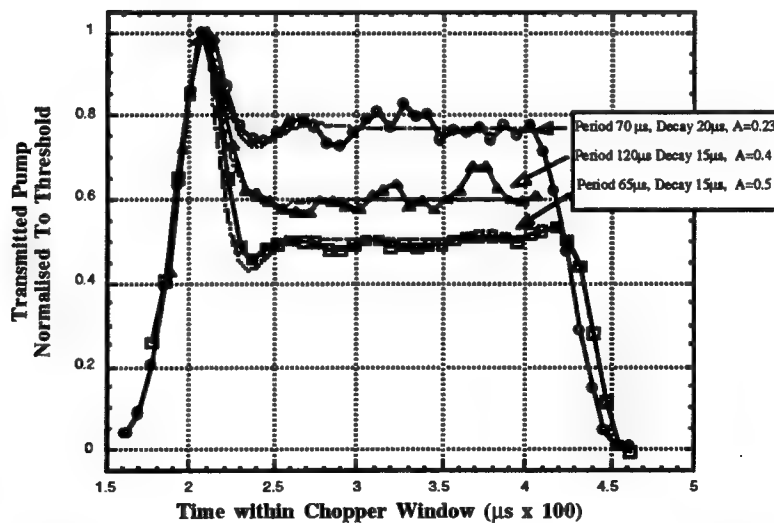


Figure 9. Plot of several fits to the experimental data, the parameters obtained are listed in the insert. Typically the period is $\sim 85 \mu$ s with a decay (1/e) $\sim 17 \mu$ s

The data and fits presented here can only be considered qualitative and preliminary. The differences between the fits is more than likely explained by noise. However the data does reveal time constants of the order of microseconds in the operation of the synchronously pumped OPO. A more careful study with more data points and less noise will probably reveal differences in timeconstants as a function of outcoupling. It is noteworthy that through the high spatial and temporal resolution of the detection systems, (micron saptial and picosecond temporal) that dynamics of over eight orders of magnitude difference have been identified. The instantaneous pump depeltion within the pump pulse typically ~ 20 ps, to the oscillation of energy between pump and signal, ~ 0.1 ms, back conversion.

I would like to gratefully acknowledge Dr Karl Koch, Dr Gerry Malone, Dr Mike DearBourn and Dr Sean Ross, all who have helped made my tour productive and less of a duty and more of a pleasure. I would like to gratefully acknowledge RDL for funding support.

References

1. Graham R. Allan, "Temporal and Spatial Characterisation of a Synchronously-Pumped Periodically-Poled Lithium Niobate Optical Parametric Oscillator."
AFOSR Summer Faculty Research Program End of Tour 96

Graham R. Allan, "Temporal and Spatial Characterisation of a Synchronously-Pumped Periodically-Poled Lithium Niobate Optical Parametric Oscillator."
AFOSR Summer Faculty Research Program End of Tour 97

STABLE CONTROLLER DESIGN FOR DEPLOYABLE PRECISION
STRUCTURES USING PERTURBATION THEORY

Mark J. Balas
Professor
Department of Aerospace Engineering

University of Colorado at Boulder
Campus Box 429
Boulder, CO 80309-0429

Final Report for:
Summer Faculty Research Program
Phillips Laboratory Site

Sponsored by:
Air Force Research Laboratory
Kirtland Air Force Base, Albuquerque, NM

September 1998

Stable Controller Design for Deployable Precision Structures Using Perturbation Theory

ROBERT J. FUENTES
MARK J. BALAS

*Center for Aerospace Structures
University of Colorado at Boulder,
Boulder, CO 80309-0429*

ABSTRACT. Perturbation methods for linear operators are commonly used in the analysis of systems that tend to deviate linearly from a given reference model. The construction of an operator with a first order perturbation is investigated and the resulting eigenvalue series is constructed. It is shown that a simplified perturbation series can be obtained for matrix operators with special structure. The given theory is applied to a reduced order model (ROM) control scenario and an algorithm for computing an $O(\varepsilon^3)$ eigenvalue approximation is described.

INTRODUCTION

Modern modeling of large dimensional physical systems primarily employs matrix methods, often via finite elements or partial differential equations. Due to implementation constraints from a control standpoint, however, an adequate model reduction is needed. Model approximations can be used to design controllers, but compensation for observation spillover may be required for stability. It has been shown in [1] that addition of a residual mode filter to the reduced order model controller can stabilize the entire system in modal coordinates. The question now becomes: What components of the spectrum were destabilized? An eigenvalue perturbation method, developed in [3], has been proposed for full rank operators with no multiplicities. In this paper, we have expanded this technique to consider multiple eigenvalues and also provide proof of the order of approximation.

Applied mathematicians, engineers and physicists frequently use perturbation methods to solve a variety of different problems. Even though solutions to intractable nonlinear differential equations can be approximated using perturbation theory, this paper's focus will only encompass linear deviations of finite-dimensional linear operators. The reader is also referred to an "in-depth" treatment of the perturbation theory of linear operators [4], where the subsequent derivations become special cases.

The following definitions lay the groundwork for the construction of an eigenvalue perturbation series. The type of structure imparted upon a matrix operator of interest is illustrated and the corresponding eigenvalue series is given. An example of the application to the control design of a spectral system is explained. A proof and the numerical algorithm for this approach are given in the appendices.

OPERATOR PERTURBATIONS

Beginning with an operator with a first order perturbation,

$$\begin{aligned} A(\varepsilon) &= A^0 + \varepsilon \Delta A \\ A^0 &\in \mathbb{R}^{L \times L} \text{ (diagonalizable, full rank)} \\ \varepsilon &\in \mathbb{C}, \Delta A \in \mathbb{R}^{L \times L} \end{aligned} \quad (1)$$

the resolvent of $A(\varepsilon)$ is defined by the following inverse:

$$R(\lambda, A(\varepsilon)) \equiv [A(\varepsilon) - \lambda I]^{-1} \quad (2)$$

and is analytic everywhere except at the eigenvalues of $A(\varepsilon)$. It has been shown in [3] that the resolvent can be written in series form,

$$R(\lambda, A(\varepsilon)) = R(\lambda, A^0) + \sum_{n=1}^{\infty} R^{(n)}(\lambda) \varepsilon^n \quad (3)$$

$$R^{(n)}(\lambda) \equiv R(\lambda, A^0) [-\Delta A R(\lambda, A^0)]^n$$

and the series is convergent if the following inequality holds.

$$0 \leq \|\Delta A R(\lambda, A^0)\| |\varepsilon| < 1$$

The spectral factorization of A^0 can be expressed as a sum of S orthogonal eigen-projections, P_l .

$$A^0 = U \bar{\Lambda} U^{-1} = \sum_{k=1}^L \bar{\lambda}_k P_k = \sum_{l=1}^S \lambda_l P_l^0 = \sum_{l=1}^S \lambda_l U E_l U^{-1} \quad (4)$$

$$E_l \equiv \begin{bmatrix} 0 & 0 & \cdot & 0 \\ 0 & I_{m \times m} & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & 0 \end{bmatrix}$$

Note that $S < N$ if any of the eigenvalues in $\bar{\Lambda}$ are repeated. The multiplicity of the l^{th} eigenvalue is computed by taking the trace of the l^{th} factor in equation (4).

$$\text{Tr}[U E_l U^{-1}] = \text{Tr}[E_l U^{-1} U] = \text{Tr}[E_l] = m_l$$

Similarly, the resolvent can be written as a sum of projections:

$$R(\lambda, A^0) = \sum_{l=1}^S \frac{P_l}{\lambda_l - \lambda} \quad (5)$$

and the reduced resolvent, $S_n(\lambda)$, is defined in equation (6) as the complement of the n^{th} projection.

$$S_n(\lambda) \equiv \sum_{\substack{l=1 \\ l \neq n}}^S \frac{P_l}{\lambda_l - \lambda} \quad (6)$$

Each of these eigen-projections can be recovered from the integration of the resolvent around a closed curve in the complex plane.

$$P_l = \frac{-1}{2\pi j} \oint_{\Gamma_l} R(\lambda, A^0) d\lambda \quad (7)$$

OPERATOR PARTITIONING

Assume that the linear operator $A(\varepsilon)$ can be partitioned in the following manner:

$$\begin{aligned}
 A(\varepsilon) &= \begin{bmatrix} A^{1,1} & \varepsilon \Delta A^{1,2} \\ \varepsilon \Delta A^{2,1} & A^{2,2} \end{bmatrix} = \begin{bmatrix} A^{1,1} & 0 \\ 0 & A^{2,2} \end{bmatrix} + \varepsilon \begin{bmatrix} 0 & \Delta A^{1,2} \\ \Delta A^{2,1} & 0 \end{bmatrix} \\
 &= A^0 + \varepsilon \Delta A \\
 A^0 &\in X_{BD} \text{ (Block Diagonal)} \\
 \Delta A &\in X_{OD} \text{ (Block Off-Diagonal)}
 \end{aligned} \tag{8}$$

The first factor, A^0 , is an element in the space of block diagonal matrices. The perturbing matrix, ΔA , is a member of the complementary block off-diagonal space. In fact, matrix multiplication of these operators can be interpreted as mappings between these spaces.

It is said that A and B are partitioned when A is a complement matrix of B (A and B do not overlap) and $A \in X_{BD}$; $B \in X_{OD}$, then the following definitions hold:

$$(1.) A^n \in X_{BD} \quad \forall n = 1, 2, \dots$$

$$(2.) B^n \in \begin{cases} X_{BD} & \text{for } n \text{ even} \\ X_{OD} & \text{for } n \text{ odd} \end{cases}$$

$$(3.) \begin{Bmatrix} A \cdot B \\ B \cdot A \end{Bmatrix} \in X_{OD}$$

EIGENVALUE PERTURBATIONS

The k^{th} projection of an operator, $A(\varepsilon)$, can be written as the corresponding eigenvalue scaling that projection.

$$A(\varepsilon) P_k(\varepsilon) = P_k(\varepsilon) A(\varepsilon) = \lambda_k(\varepsilon) P_k(\varepsilon) \tag{9}$$

$P_k(\varepsilon)$ is determined from integrating the resolvent, $R(\lambda, A(\varepsilon))$, around the k^{th} eigenvalue and results in an infinite series in ε .

$$\begin{aligned}
 P_k(\varepsilon) &= \frac{-1}{2\pi j} \oint_{\Gamma_k} R(\lambda, A(\varepsilon)) d\lambda = \frac{-1}{2\pi j} \oint_{\Gamma_k} R(\lambda, A^0) d\lambda + \frac{-1}{2\pi j} \varepsilon^n \int \sum_{n=1}^{\infty} R^{(n)}(\lambda) d\lambda \\
 &= P_k^0 + \sum_{n=1}^{\infty} P_k^{(n)} \varepsilon^n
 \end{aligned} \tag{10}$$

By taking the trace of equation (9), the k^{th} eigenvalue series is recovered:

$$\begin{aligned}
 \lambda_k(\varepsilon) &= \frac{1}{m_k} \text{Tr}[A(\varepsilon) P_k(\varepsilon)] \\
 &= \lambda_k^0 + \sum_{n=1}^{\infty} \lambda_k^{(n)} \varepsilon^n
 \end{aligned} \tag{11}$$

where λ^0 represents an eigenvalue of A^0 . If $A(\varepsilon)$ is partitioned into the form given in (8), the previous equation is simplified to an even series.

$$\lambda_k(\varepsilon) = \lambda_k^0 + \lambda_k^{(2)} \varepsilon^2 + \lambda_k^{(4)} \varepsilon^4 + \dots = \lambda_k^0 + \sum_{n=1}^{\infty} \lambda_k^{(2n)} \varepsilon^{2n} \quad (12)$$

Proof of this result is given in Appendix A.

DETERMINATION OF INSTABILITY IN LINEARLY PERTURBED SYSTEMS

At this point, the application of this method to the state space control approach given in [1], [2], and [5] is considered. Given a linear plant governed by the following vector differential equation:

$$\begin{aligned} \dot{\underline{x}} &= A \underline{x} + B \underline{u} \\ \underline{y} &= C \underline{x} + D \underline{u} \end{aligned} \quad (13)$$

$\underline{x} \equiv$ plant state vector

$\underline{u} \equiv$ control input vector

$\underline{y} \equiv$ output observation vector

the objective is to stabilize (13) through the use of M actuator inputs and P sensor outputs. Matrix operators B and C will have the finite rank of M and P, respectively, with the linear differential operator, A , possibly of infinite dimension.

Using the ROM (reduced order model) methodology, it is possible to transform the plant to a modal coordinate system and separate (13) into two state equations. Assuming the feed-through term, D , is zero:

$$\begin{aligned} \dot{\underline{x}}_N &= A_N \underline{x}_N + B_N \underline{u} \\ \underline{y}_N &= C_N \underline{x}_N \end{aligned} \quad (14a)$$

$$\begin{aligned} \dot{\underline{x}}_R &= A_R \underline{x}_R + B_R \underline{u} \\ \underline{y}_R &= C_R \underline{x}_R \\ \underline{y} &= \underline{y}_N + \underline{y}_R \end{aligned} \quad (14b)$$

The preliminary control design of the entire plant is based upon the dynamics of (14a). A finite number of modes are chosen for A_N , which is an operator of rank N. Equation (14b) consists of modes that are not of immediate consideration and are open loop stable. ($\text{Re}(\lambda(A_R)) < 0$) The use of state estimation and N state feedback results in the system of equations in (15).

$$\begin{aligned} \dot{\hat{\underline{x}}}_N &= A_N \hat{\underline{x}}_N + B_N \underline{u} + K_N (\underline{y} - \hat{\underline{y}}_N) \\ \hat{\underline{y}}_N &= C_N \hat{\underline{x}}_N \\ \underline{u} &= G_N \hat{\underline{x}}_N \end{aligned} \quad (15)$$

Defining the error between estimated and actual states by a vector \underline{e}_N , the differential equation governing error progression becomes the difference between state equations:

$$\underline{e}_N = \dot{\hat{x}}_N - \underline{x}_N \quad (16)$$

$$\dot{\underline{e}}_N = (A_N - K_N C_N) \underline{e}_N + K_N C_R \underline{x}_R$$

If controllability and observability requirements are met for the system, arbitrary decay rates for the designed subsystem can be achieved by choosing the appropriate estimator gain, K_N , and feedback gain, G_N , matrices for the desired eigenvalues of $A-KC$ and $A+BG$.

Writing equations (14a), (14b), and (16) in matrix form, the probable causes of instability in the closed loop system can be clearly determined.

$$\begin{bmatrix} \dot{\underline{x}}_N \\ \dot{\underline{e}}_N \\ \dot{\underline{x}}_R \end{bmatrix} = \underbrace{\begin{bmatrix} A_N + B_N G_N & B_N G_N & 0 \\ 0 & A_N - K_N C_N & K_N C_R \\ \mathbf{B}_R \mathbf{G}_N & \mathbf{B}_R \mathbf{G}_N & A_R \end{bmatrix}}_{A_c} \begin{bmatrix} \underline{x}_N \\ \underline{e}_N \\ \underline{x}_R \end{bmatrix} \quad (17)$$

Cross-coupling terms in (17), illustrated in bold, characterize the introduction of estimator feedback which can destabilize the residual state equation. One option to remedy this problem is given in [1], in which a residual mode filter (RMF) is developed to compensate for eigenvalues that are drawn unstable. However, knowledge of those open loop eigenvalues contained in A_R is needed in order to implement the method. Previous work by Gooyabadi [3] offers a limited perturbation solution to this problem.

Partitioning equation (17) according to the method given in section (2), the matrix A_c can be considered to be the sum of a stable, linear operator, A^0 , and a perturbing operator, ΔA .

$$A(\varepsilon) \equiv \begin{bmatrix} A_N + B_N G_N & B_N G_N & 0 \\ 0 & A_N - K_N C_N & 0 \\ 0 & 0 & A_R \end{bmatrix} + \varepsilon \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & K_N C_R \\ B_R G_N & B_R G_N & 0 \end{bmatrix} \quad (18)$$

$$= A^0 + \varepsilon \Delta A$$

Our objective is to find out which eigenvalues of A^0 are driven unstable by ΔA through the computation of the first few non-zero terms of the perturbation series (12). Taking $\varepsilon = 1$ would give equation (17), but convergence of the eigenvalue series solely rests upon the condition that $\|\Delta A R(\lambda, A^0)\| < 1$. Therefore, it is possible that (12) is divergent for arbitrary $\|\Delta A R(\lambda, A^0)\| \neq 1$ and may have asymptotic convergence.

An example of implementation of this perturbation scheme is given for a simple Euler-Bernoulli beam with pinned ends:

$$m z_{,tt}(x,t) - 2\xi z_{,txx}(x,t) + E \cdot I z_{,xxxx}(x,t) = u(t) \delta(x - 0.1 \cdot l) \quad (19)$$

$$\text{B.C.:} \quad \begin{aligned} z(0, t) &= z(l, t) = 0 \\ z_{,xx}(0, t) &= z_{,xx}(l, t) = 0 \end{aligned}$$

$$y(t) = \int_0^l z(x,t) \delta(x - 0.9 \cdot l) dx = z(0.9 \cdot l, t)$$

All of the constants in (19) are normalized with the exception of the damping ratio. ($|\xi| = 5.0 \cdot 10^{-6}$) Actuation of the beam occurs at the point $x = 0.1$. Observations of the beam

displacement, $y(t)$, are taken at $x = 0.9$. Using the method of eigenfunction expansions, a temporal differential equation can be constructed:

$$\ddot{z}_k(t) = -(k\pi)^4 z_k(t) - 2\xi(k\pi)^2 \dot{z}_k(t) + \phi_k(0.1) u(t) \quad (20)$$

The spatial eigenfunctions, ϕ_k , transform the input, $u(t)$, into the modal system of equations given in (20). Control is introduced to dampen the first few modes, where $k = 1, 2$, and 3 . There exist an infinite number of residual modes ($k \geq 4$), but for computational purposes only the next 12 are considered. Putting the equations derived from (13) in the form given in (18), we can calculate an estimate of the residual eigenvalue perturbation resulting from the control design:

$$\lambda_k(\varepsilon)|_{\varepsilon=1} \cong \lambda_k^0 + \lambda_k^{(2)} \quad (21)$$

The results of this $O(\varepsilon^3)$ approximation are illustrated below in Figure 1. from this graph, it can be determined that modes $n = 5, 7$ are driven unstable. Using these modes in an RMF control design, exponential stability for the entire system can be achieved.

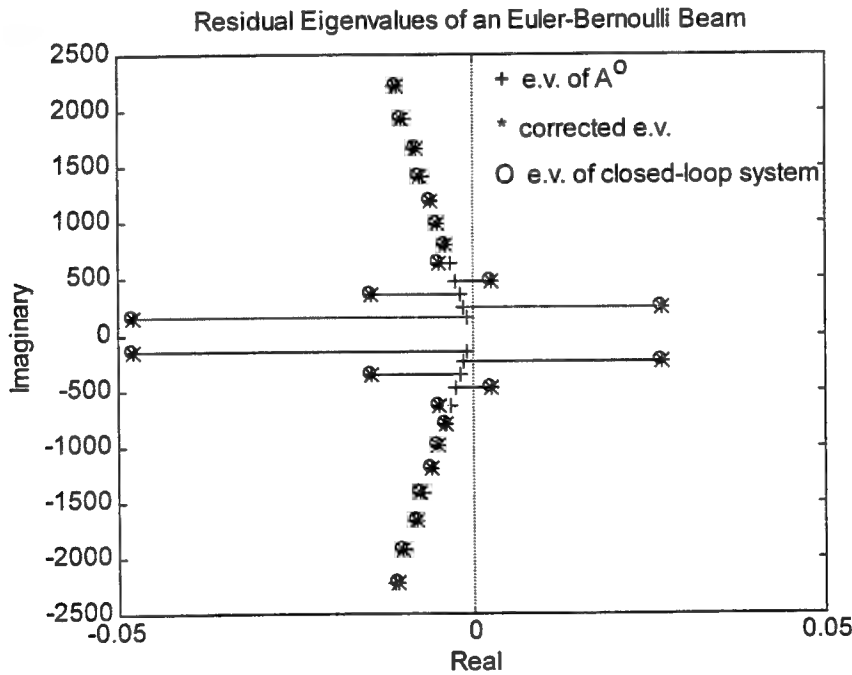


FIGURE 1.

CONCLUSIONS

A generalized approach to the perturbation method outlined in [3] has been developed. Given a block diagonal linear operator with a first order perturbation in partitioned form, it has been shown that there exists a corresponding even perturbation series. An algorithm is outlined to find the $O(\varepsilon^3)$ correction for block diagonal operators

containing eigenvalue multiplicities. Application of this technique to a simple structural model illustrates that this method is a viable approach to ROM control design. This procedure is shown to be successful in other structural problems with no repeated eigenvalues using finite element models [3].

ACKNOWLEDGMENTS

This research was supported by the Air Force Office of Scientific Research Summer Research Program and the National Science Foundation. Any opinions, findings, and conclusions stated in this publication are those of the authors and do not necessarily reflect the views of AFOSR or NSF.

REFERENCES

1. M. Balas, Finite-Dimensional Controllers for Linear Distributed Parameter Systems: Exponential Stability Using Residual Mode Filters, *J. Math. Anal. Appl.*, Vol. 133, No.2, 1988.
2. M. Balas et. al., Low Order Control of Large Aerospace Structures Using Residual Mode Filters, CU-CSSC-88-05, 1988.
3. A. Gooyabadi, Doctoral Thesis, University of Colorado at Boulder, Boulder, CO, 1992.
4. T. Kato, Perturbation Theory for Linear Operators, Springer-Verlag, New York, 1980 Reprint.
5. B. Reisenauer and M. Balas, Master's Thesis, CSI Compensation for Reduced-Order Model Based Control of a Flexible Robot Manipulator, 1990.

IMAGE RECOVERY USING PHASE DIVERSITY

Neb Duric
Professor
Department of Physics and Astronomy

University of New Mexico
800 Yale Bld NE
Albuquerque, NM 87131

Final report for:
Summer Faculty Research Program
US Air Force Research Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force base, DC

and
US Air Force Research Laboratory

October, 1998

IMAGE RECOVERY USING PHASE DIVERSITY

Neb Duric
Professor
Department of Physics and Astronomy

ABSTRACT

The angular resolution of ground-based astronomical imaging is limited by turbulence in the earth's atmosphere. This turbulence varies both spatially and temporally so that it acts as a rapidly fluctuating aberrator across the aperture of a telescope. We report on a laboratory experiment aimed at characterizing the ability of the phase diversity technique to correct for such aberrations. The object being imaged and the aberrations are both generated in the laboratory. An astronomical object is simulated with a laser/pinhole combination and a series of lenses that effectively place the object at infinity. The turbulence is simulated with various stationary and time-varying aberrators such as plastic films and locally heated air. Two short-exposure images of the simulated object are taken with a CCD camera. One conventional image (in focus) and one diversity image, which is defocused by a fixed amount, where both are degraded by the same aberrations. Since the shape of an image located at an arbitrary position with respect to the focal plane is a strong function of the aberrations, the pair of images recorded with this phase diversity imaging technique, together with a maximum-likelihood estimate is used to solve for both the aberrations and the object that most likely represents the recorded images.

The starting point of the experiment was to study a point-like object (a star) in the presence of stationary aberrations in the limit of strong signal statistics. The ultimate goal is to test the phase diversity technique in the presence of strongly time varying aberrations in the faint signal limit and with various combinations of the D/r_0 ratio where D is the instrument aperture and r_0 is the coherence length of the turbulent air.

The results of image recovery in the presence of stationary aberrations are presented. The role of detector noise and dynamic range in limiting the recovery are also discussed.

1 Introduction

The Earth's atmosphere acts as an aberrator in which the aberrations vary spatially *and* temporally. For imaging objects at infinity one can consider the atmosphere as disturbing a wavefront which is initially plane-parallel, that is, surfaces of constant phases are planes oriented at right angles to the direction of propagation. As the plane wave travels through the Earth's atmosphere, different parts of the wavefront undergo different amounts of phase retardation resulting from the fact that different patches of the Earth's atmosphere have different indices of refraction. By the time the wavefront reaches the ground-based detector it is no longer plane-parallel. It has an overall tilt and the surfaces of constant phase are now corrugated. The corrugations represent phase variations (aberrations) imposed by the earth's atmosphere. For a wavefront incident on a circular aperture such as the aperture of a telescope, the aberrations can be expanded in a convenient series of functions called "Zernike Polynomials". Thus, any aberration can be, in principle, characterized by an appropriate set of Zernike polynomials. Any technique that can determine the coefficients of the Zernike polynomials can yield a characterization of the aberration and therefore can "undo" the aberration through an appropriate deconvolution.

The phase diversity technique takes advantage of the fact that the morphology of an image located at an arbitrary location with respect to the focal plane is a strong function of the aberrations that are present. Thus, the appearance of a defocused image, relative to an in-focus image, depends very much on the aberrations that the two images share. We now present results of an experiment aimed at recovering images using the "Gonsalves Error Metric" approach (first proposed by Gonsalves and Childrow, 1979).

2 Experimental Setup

We have set up a phase diversity camera and performed experiments aimed at recovering high-resolution images (see figure below). A system of lenses on an optical bench is used to focus images through the entrance pupil of the phase diversity camera. The camera was constructed by Lexitec Inc.

In order to simulate astronomical observations, we used a re-imaging setup whereby all focused images appear as if they are at infinity relative

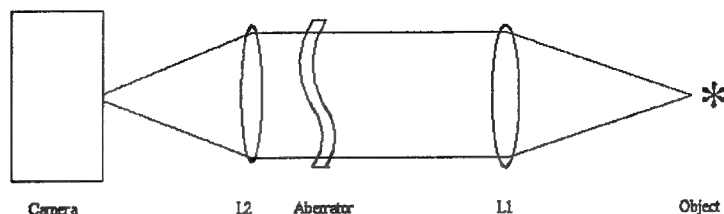


Figure 1: The Experimental Setup

to the camera. This is achieved by placing the object to be imaged at the focal point of the collimating lens, L1. An aberrator, is placed near the second converging lens, L2. In doing so we mimic the atmosphere as being a relatively local aberrator relative to the distant source of light such as a geostationary satellite or an astronomical object.

2.1 Correcting Defocus

With this experimental setup we have initiated a stepped approach toward the ultimate goal of recovering images of astronomical objects. The first step is illustrated in figure 2 which shows the recovery of an image in the presence of aberrations that are dominated by a defocus term. The top left and top right frames show the in-focus and out-of-focus aberrated images respectively. The OTF corresponding to the calculated aberrations is shown on the bottom left. The recovered image (obtained by deconvolving the PSF from the top left image) is shown on the bottom right. The full width at half maximum (FWHM) of the recovered image was similar to that of the original image. The image recovery therefore appears to be able to re-sharpen images that have been intentionally de-focused.

2.2 Correcting Complex Aberrations

Figure 3 shows the recovery of an image in the presence of more complex aberrations.

In this case, a rigid piece of flawed plastic was used to distort an image that is initially roughly a point source (in fact a HeNe laser spot). The error

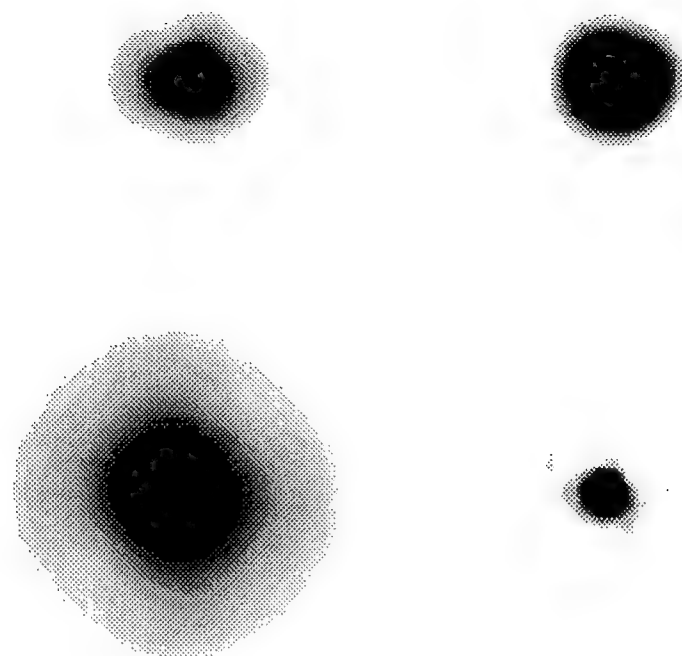


Figure 2:

minimization technique was used in combination with 19 Zernike polynomials (the piston and tip and tilt terms were set to zero) in the image reconstruction process. The aberrated image is shown in the top left corner of figure 3. The defocused aberrated image is shown in the top right. The OTF is shown in the lower left corner, as before. The recovered image is characterized by a bright central core surrounded by an extended distribution of relatively faint residuals.

Since these images are dominated by Poisson noise it is expected that that the brightest portions of the images will be recovered with the greatest accuracy. Consequently the quality of the recovery declines with falling light intensity. This effect, combined with the limited number of Zernike polynomials that can be used, appears to account for the diffuse plateau of residuals.

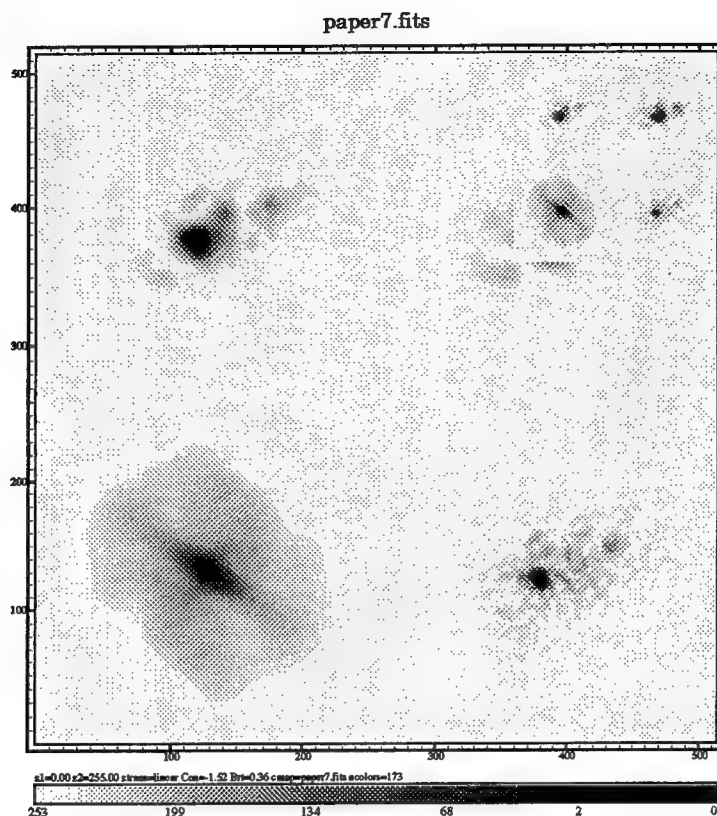


Figure 3:

The core has the expected FWHM indicating that the recovery was generally successful.

2.3 Image Alignment

It was found that alignment of the images played a crucial role in the determining the quality of the recovered image. We tried a number of different approaches in order to align the in-focus and out-of-focus images. Interactive centering with a cursor was limited to an accuracy of no better than a pixel. This proved insufficient for reliable (reproducible) image recovery. Our second approach was to fit Gaussians to the input images. This worked well

for sub-pixel alignment but only with simple point-like images. The best method turned out to a cross-correlation method. A code was written in IDL to cross-correlate the in-focus and out-of-focus images yielding a cross-correlation image such as the one shown in figure 4. The cross-correlation

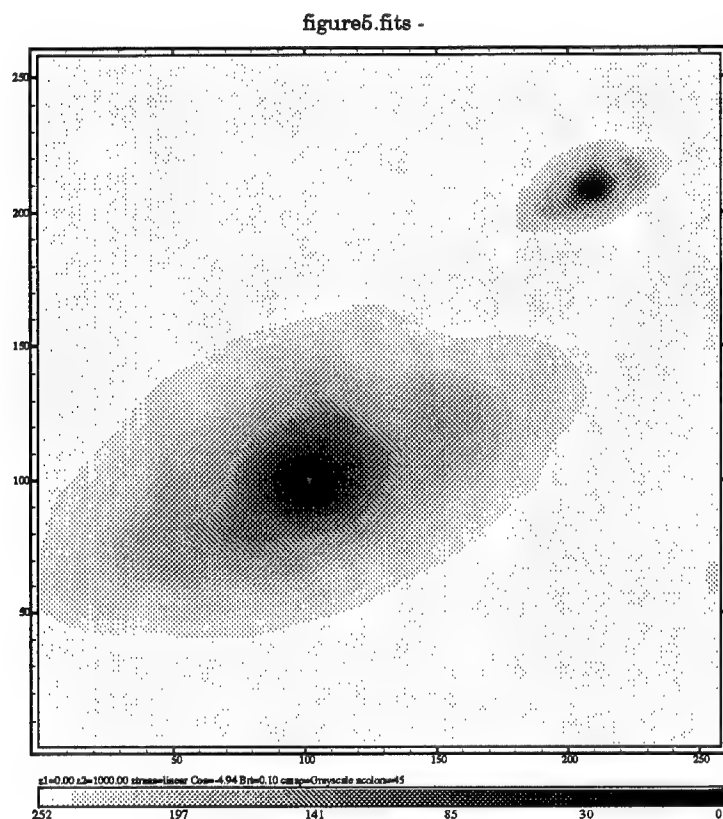


Figure 4:

image has the useful feature that it has a strongly peaked central core even when the input images are complex. The location of the core relative to the center of the in-focus image yields the offset between the in-focus and out-of-focus images. Further Gauss-fitting of the core of the cross-correlation function leads to sub-pixel determination of the offset and accurate and reproducible alignment of the images.

3 Limitations of the Image Reconstruction

Our experiments are limited by two major factors, the finite well depth of the CCD camera and the Poisson statistics. The two effects are coupled for the following reason. Since the in-focus and out-of-focus images are obtained simultaneously an in-focus image that is just under the saturation threshold will have an optimal signal to noise distribution as governed by the Poisson statistics. However, the out-of-focus image (which is slaved to the in-focus image with regard to exposure time) will have a less optimal signal to noise distribution because its light is spread out more. Consequently the recovery process is limited by the inferior signal to noise characteristics of the defocused image. Thus, even in the strong signal limit, the image recovery is limited. Of course, co-adding of images can remedy this problem but only in the case of static, repeatable aberrations. This will, in general, not be possible for time-varying aberrations. Thus we see that on the faint end the recovery is limited by the combination of instrumental and Poisson noise while at the strong signal end the recovery is limited by the combination of finite well-depth and Poisson noise. The latter can be somewhat alleviated through the use of CCDs that are configured for greater well depth.

4 Future Work

The use of more complex stationary aberrators will be used to determine the effectiveness of the phase diversity technique in recovering images of known morphology. This will be achieved through the experimental setup shown above, where collimated light from a known image will be passed through an aberrator before being focused onto the phase diversity camera. By varying the light level of the received signal and by varying the spectrum of aberrations the limitations of the phase diversity method will be explored.

The next step is to introduce time-varying aberrations into the imaging system to better replicate the turbulent nature of the earth's atmosphere. The atmosphere can also be thought of as a phase screen but one which changes its character on time scales of $\tau_0 \approx 1 - 100$ ms. The phase diversity technique is therefore expected to work on images recorded over time scales shorter than the coherence time of the atmosphere. Time varying aberrations can be produced in the lab by using rotating pieces of aberrated glass or by

generating turbulent motion of air with a localized heat source. We intend to use both kinds of aberrators in our experiments.

The ultimate success of the image recovery will depend on two competing effects, the integration time, τ , used to record the signal, and the coherence time, τ_0 , of the aberrations. Thus, it is desirable to use as large a τ as possible, in situations where the light level is low but at the same time, $\tau < \tau_0$ is the necessary requirement for optimal use of the phase diversity technique. For $\tau \gg \tau_0$, the phase diversity technique fails. There is therefore a grey area, defined by $\tau \approx \tau_0$, where the phase diversity technique does not work optimally but may nevertheless provide significant corrections to the aberrations. Our goal is to fully explore and document this regime because of its relevance to image recovery at low light levels.

5 References

- Gonsalves, R.A. and Childlow, R. 1979, SPIE, 207, 32.
Gonsalves, R.A. 1982, *Optical Engineering*, 829, 21.
Restaino, S. 1992, Appl. Opt. 31, 7442.

THEORY OF HYDROGEN IN SiO_2

A. H. Edwards
Associate Professor
Department of Electrical Engineering

University of North Carolina at Charlotte
Charlotte, NC 28223

Final Report for:
Summer Faculty Research Program
Air Force Research Laboratory
VSSE

Sponsored by:
AirForce Office of Scientific Research
Bolling Air Force Base, DC

and

Phillips Laboratory

October 1998

THEORY OF HYCROGEN IN SIO2

A. H. Edwards
Associate Professor
Department of Electrical Engineering

Abstract

We report a variety of quantum chemical and electrostatic calculations that pertain to the capture of electrons by positively charged hydrogenic species in SiO_2 and at the Si-SiO_2 interface. We have calculated *ab initio* optical absorption energies for the Frank-Condon transitions. We place our results in the context of a previously developed model that relies on vibrational excitation. The estimate for excitation energies depends significantly on the approximation used for the excited states. However, all Frank-Condon transitions imply the need for important refinement of the original model. As an adjunct to these calculations, we have performed the first *ab initio* calculation of the proton *at* the Si-SiO_2 interface. We find that the proton is the stable form of positive charge in the MOS system. That is, a proton in SiO_2 is predicted to be more stable than a hole in the silicon valence band and a neutral hydrogen atom in SiO_2 . This is in agreement with previous experiment on unirradiated SiO_2 .

1 Introduction

Over the last several years several groups have explored the physics of positively charged hydrogenic species, that we will call protons, in SiO_2 . Vanheusden and Stesmans were the first to observe charging in buried oxides after annealing in forming gas at high temperatures[1]. This effect has been studied experimentally by Vanheusden and coworkers [2,3], and by Stahlbush [4], and theoretically by Edwards and coworkers [5], Ferreira and coworkers [6] and, most recently, by Karna and coworkers [7]. All theoretical studies agree on the structure of the proton. We illustrate the conformation in Fig. 1 that depicts the cluster used for most of the quantum chemical calculations presented here.

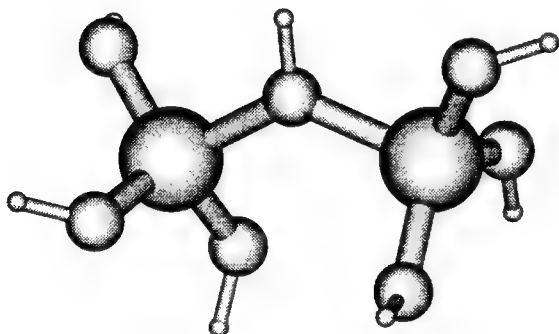


Figure 1: Cluster used to simulate a proton in bulk SiO_2 . Small balls are hydrogen atoms, middle-sized balls are oxygen atoms and large balls are silicon atoms. Cluster shown in equilibrium geometry.

As is obvious from the figure, the name proton is only a convenience. Rather than an isolated, bare proton, this moiety is a three fold coordinated oxygen that is only stable in the positive charge state. There are several interesting aspects of this defect. One is that it is stable near the Si- SiO_2 interface. Vanheusden and coworkers have, under the influence of an electric field, swept protons to this interface causing the flow of electron current in the substrate [8]. Thus, even when there are large numbers of electrons in the silicon conduction band, the protons are not annihilated. We note that this is not the case in irradiated oxides [9]. There, protons are annihilated and this annihilation is correlated with an increase in interface states. Another interesting feature of this defect is that under certain circumstances this positively charged defect will not capture electrons at low temperatures [2]. We have previously articulated a general model that would lead

naturally to this observation [2]. One obvious prediction is that for 5.5 eV photons, there should be significant charge injection into the oxide from the substrate, if we assume that the positive charge in the oxide generates large fields. We have performed an electrostatics calculation in two different approximations that contradicts that assumption. The two approximations agree very well with one another. Here we present a detailed calculation of ground- and excited- state potential surfaces for proton dissociation. We have calculated excited states using both single reference configuration interaction (CI) and Restricted Open-shell Hartree-Fock (ROHF) theory. These results agree to a remarkable extent and, if true, preclude the simple model put forward in Ref. [3]. The balance of the report is organized as follows: In Sect. 2, we discuss the methods of calculation for the quantum chemical calculations, the vibrational analysis, and the electrostatic calculations. In Sect. 3, we present the cluster calculations that simulate the bulk. In Sect. 4 we present calculations on the proton at the Si-SiO₂ interface. In Sect. 5 we present, briefly, the electrostatic calculations that speak to the photo-injection of electrons into the oxide. Finally, in Sect. 6 we discuss our results in light of the experiments in unirradiated oxides.

2 Methods of calculation

We used two molecular orbital packages, GAMESS [10] and GAMESS-UK (see acknowledgments) in all of the current calculations. As mentioned in Sect. 1, we used the cluster in Fig. 1 for all of the calculations on the bulk proton undertaken during the summer of 1998. In Sect. 3 we report the results of a study of basis sets and level of CI. As a result of this study, in most of the calculations, we use the 6-31** basis. Six Gaussian primitive functions are used to represent the set of core states, while the valence set is split: an inner atomic orbital with smaller radius is represented by three Gaussian functions, and an outer atomic orbital with larger radius is represented by one Gaussian. This splitting allows valence orbitals to change shape as a function of their environments. The 6-31G** basis also contains polarization functions (*d*- functions on the heavy atoms and *p*- functions on the hydrogen atoms). We have augmented this basis with diffuse *sp*- orbitals on the heavy atoms. We will discuss in detail this choice in Sect. 3. To calibrate large cluster results, we have performed other calculations in which we used the Stevens-Basch-Krauss effective core potential [11] with its associated basis set. To review briefly, an ECP is the molecular equivalent of the solid state pseudo-potential. All core electrons are replaced by a potential that depends on angular momentum. The valence electrons interact with this ECP as a normal one-electron operator and with each other through the normal Coulomb repulsion. This potential function reproduces the one-electron valence eigen values obtained in a normal all-electron calculation. The valence basis set

is a 3-1 split valence set, similar to those in the 6-31G** set. We have also augmented the basis with *d*-polarization functions and *sp*-diffuse functions on the heavy atoms, and *p*-polarization elements on the hydrogen atoms, again mirroring the 6-31G** basis. We have studied the effects of electron correlation on the ground states using both second order Møller-Plesset perturbative CI [12], and single reference- CI including all single and double excitations. For excited states, we used the single reference CI and ROHF theory.

To calculate vibrational spectra, we used GAMESS to obtain all modes in the harmonic approximation. To include anharmonic effects in the stretching and wagging modes, we used our own finite element code based on cubic spline elements. We used the same code to obtain continuum wave functions from the excited state potential curves. We used a cubic spline interpolation to obtain additional values of the potential surface required for the fine gridding necessary in the continuum calculations, and for the equi-spaced gridding used for the vibrational levels of the ground electronic state.

To address the likelihood of photo-injection, we have used the depletion approximation applied to an MOS capacitor. This is a standard approximation that gives remarkably good results. It is discussed fully in a variety of device physics texts (See, for example, [13]). We have compared the results of depletion approximation with those obtained using our own one dimensional device modeling code based on linear finite elements.

3 Protons in bulk SiO₂

3.1 Basis set and electron correlation study

To assess the relevance of the model we put forward previously [2], we need reliable ground- and excited state- potential surfaces. To assess reliability, we have performed a convergence study of the adiabatic excitation energy at the ground-state equilibrium geometry of the cluster shown in Fig. 1. We used the fully relaxed geometry for the positive charge state that is detailed in Table 1

We have shown the geometrical parameters for both the proton and the neutral cluster without the central hydrogen atom to give an appreciation of the size of the relaxation. In Table 2, we give the excitation energies as a function of the number of electrons correlated and of the number of virtual orbitals (orbitals not occupied in the ground state). Using the 6-31G** basis.

	Proton	SiO ₂
R_{Si-O^*}	1.768	1.617
R_{Si-O}	1.594	1.625
$\angle Si-O^*-Si$	128.9	148.7
$\angle O^*-Si-O$	102.8	109.5
R_{O^*-H}	0.95	

Table 1: Equilibrium geometry for the cluster shown in Fig. 1 in the positive (with central hydrogen) and neutral (without central hydrogen) charge states. Both geometries were obtained with the 6-31G** basis and all atoms were allowed to relax without constraint.

$N_{val.}$	$N_{vir.}$	$E_{g.s.}$ (a. u.)	$E_{ex.s}$ (a. u.)	$\Delta E(eV)$
6	6	-0.044	11.853	11.90
12	6	-0.092	11.614	11.49
6	12	-0.124	11.795	11.70
12	12	-0.275	11.245	11.31
24	12	-0.908	10.777	11.47
13	24	-1.181	10.560	11.53

Table 2: Ground and Excited states for various CI calculations (6-31G**): Equilibrium geometry. All energies relative to HF ground state and in eV

The 13/24 calculation represents something close to the limits of our calculational capacity. Interestingly, it seems that including more virtual orbitals is more important than including more electrons in the correlation. For all CI calculations on the bulk proton, we correlated 26 electrons into 24 virtual orbitals. In recent multi-reference CI calculations, Pacchioni reported that inclusion of diffuse functions is requisite to obtain good agreement with experimental energies. We have studied the effects of their inclusion in these calculations. In Table 3, we present our results.

Here we see a dramatic difference from inclusion of diffuse *sp* functions on the heavy atoms, and a much smaller effect from inclusion of diffuse *s*- functions on the hydrogen atoms. We should also point out that the inclusion of hydrogenic diffuse functions led to greater difficulty in achieving self-consistency. For these reasons, in all of the calculations of excitation energies reported below, we have included diffuse *sp*- functions on the heavy atoms.

Basis	$E_{g.s.}(HF)$	$E_{gs}(CI)$	$E_{ex.s}$	ΔE
6-31G**: (<i>sp</i>)	-0.5	-0.965	9.86	10.825
6-31G**: (<i>sp</i>) (<i>s</i>)	-0.52	-0.753	10.05	10.80

Table 3: Ground and Excited states for various CI calculations. For all cases, $N_{val} = 13$, $N_{vir}=24$: Equilibrium geometry. All energies relative to HF ground state (6-31G**) and in eV

3.2 Potential surfaces

In Fig. 2 we show the potential surfaces for the ground and first excited electronic states of the proton. The surfaces in Fig.2 (a) were obtained from the 13/24 CI calculations, while those shown in Fig. 2 (b) were obtained from restricted Hartree Fock singlet and ROHF calculations. In the latter case, both potential surfaces were obtained self-consistently. In these calculations the equilibrium geometry is that of the proton. The variable is R_{O^*-H} . All other geometrical parameters have been held fixed.

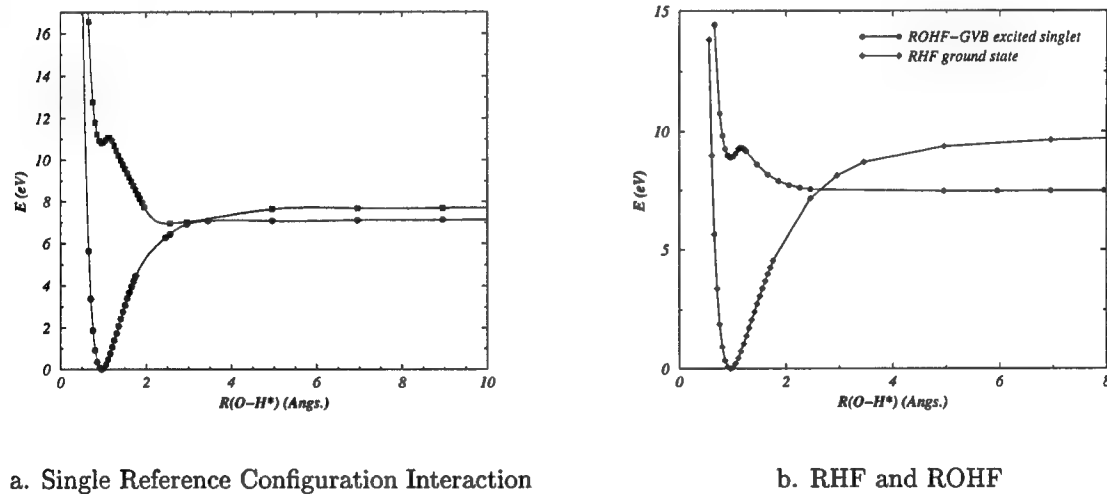


Figure 2: Potential surfaces for the ground and first-excited states of the cluster in Fig. 1.

Several features bear noting. First, the excitation energies are much larger than the experimental value. Even with a distribution of photon energies of up to 6.2 eV [14], simple thermal excitation at room temperature could not lead to optical excitation onto a repulsive excited potential surface. It may be argued that these CI calculations are only single reference, and thus might not give the best estimate of the excitation energy. However, based on our convergence study, and on previous experience with multi-reference CI calculations, the changes might be on the order of 0.5-1.0 eV . Furthermore, the SCF Hartree-Fock calculation, shown in Fig. 2 (b) gives a similarly large excitation

energy of 8.9 eV. All of these energies are far above the experimental photon energy. The calculations presented here do not support the simple model put forward in Ref. [2]. We will discuss below some possible sources of error, and some simple modifications to the model that might account for the this large disagreement.

Another equally interesting feature of this calculation is that the lowest excited state potential surfaces are not purely repulsive. Rather, they display a minimum at the ground state equilibrium geometry. There are several ways to describe this sort of behavior. In terms of molecular physics, these would be a Rydberg states and would have the same meaning as the Rydberg series in atomic spectra. That is, they reflect the attraction between the excited electron and the hole it leaves behind. In terms of solid state physics these are bound excitons. In either case, the ground state equilibrium persists because the state into which the electron is excited does not have strong antibonding character. Rather, the electronic state is rather diffuse and is analogous to a large hydrogenic orbit. We note that the same minimum would be observed if we promoted an electron into a conduction band-like orbital that has little defect character. In the case presented here, the distinction is hard to make between a bound exciton and the conduction band. We present the orbital composition of the LUMO in Table 4, along with the LUMO from the neutral cluster.

	Proton	neutral cluster	SiO ₂
<i>Si</i> - <i>s</i>	0.537	0.557	0.14
<i>Si</i> - <i>p</i>	0.02	0.05	
<i>O</i> * - <i>s</i>	0.163	0.073	
<i>O</i> - <i>s</i>	0.258	0.296	0.19
<i>H</i> * - <i>s</i>	0.004		

Table 4: Orbital decomposition of the LUMO for the proton, the neutral cluster, and SiO₂.

The oxygen contributions are almost purely from the diffuse basis elements, while the Si *s*- contributions are a mixture. We should note that in the decomposition of the SiO₂ conduction band the sum of contributions is not normalized because of a considerable contribution from plane waves. The contributions from Si and O *s*- functions dominate all three cases. In addition, the predicted excitation energy is larger than the SiO₂ band gap. Finally, the first and second excited state potential surfaces are very close to each other leading one to conclude that these states are close to the continuum limit, that is, very close to the conduction band. Whether this is conduction band admixture or bound excitons is not actually crucial to the current discussion. Rather, we are interested in whether this minimum could play a role in the dissociation process, assuming that an electron

is trapped onto that excited potential surface. While the numerical agreement is between this well depth and the observed activation energy is tantalizing, we must remember that hydrogen is light, so that kinetic energy may dominate over any tendency to localize. To explore the extent to which the minimum actually traps the hydrogen, we calculated wave functions for hydrogen motion using the one-dimensional excited potential surfaces obtained from the ROHF-GVB and CI methods. We calculated the wave functions and energies using an 800 point grid. The results are shown in Fig. 3.

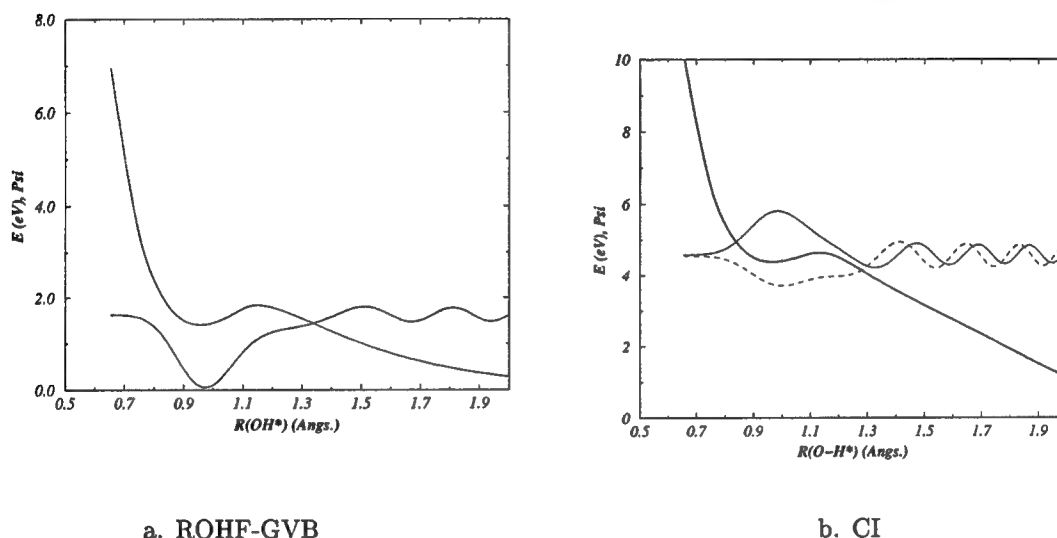


Figure 3: Resonance state of the Rydberg state-induced local minimum of the excited state potential surface.

While both surfaces have resonances, only for the excited state singlet calculation using the ROHF-GVB formalism gives strong localization. The Rydberg well generated in that case was ≈ 0.2 eV deeper than for the single reference CI results. This should have two important consequences. First, This state would give the largest vibronic probability for optical absorption. Second, because the resonance is so strong, the hydrogen will spend a considerable amount of time near the oxygen atom, leading to a possible re-emission of the electron to the SiO_2 valence band. We can see, however, that our theoretical predictions are very sensitive to the shape of the potential surface, which, in turn, depends on the method of calculation.

The final feature in Fig. 2 worth noting is that at 2.5 Å the potential surfaces appear to cross. Because the cluster has only C_1 symmetry all potential surfaces are of the same symmetry type and crossings are forbidden in the adiabatic approximation. However, with the curves approaching so closely, the adiabatic approximation may break down, at which point there will be finite probabilities that either dissociation limit can be reached from either initial curve. That is, starting from the

ground state curve, the system may dissociate into either final states depending on the velocity of the hydrogen atom as it approaches the crossing point.

Because of the large calculated excitation energies, we explored other possible pathways to dissociation that could exhibit activation at modest temperatures. We have considered dissociation through the two rocking motions- in-plane and out-of-plane. To this end, we have calculated ground state potential surfaces for hydrogen motion normal to and coincident with the Si-O-Si plane. For each case all parameters were fixed except for the appropriate angle. From these potential surfaces, we have calculated vibrational spectra, using our own finite element program. This program yields exact numerical wave functions and energy values, from which we have calculated the parameters ω_e , $\omega_e x_e$, and $\omega_e y_e$, from the usual perturbation approximation, by fitting the spectrum to the form

$$E_n = \hbar\omega_e(n + \frac{1}{2}) + \omega_e x_e(n + \frac{1}{2})^2 + \omega_e y_e(n + \frac{1}{2})^3$$

The results are given in Tables 5 and 6

Mode	$h\nu_0$	$h\nu_1$	$h\nu_2$	$h\nu_3$	$h\nu_4$
in-plane	0.075 (606)	0.226 (1820)	0.377 (3038)	0.528 (4260)	0.680 (4260)
out-of-plane	0.0004 (3.056)	0.001 (9)	0.002 (15)	0.003 (22)	0.003 (28)

Table 5: Exact energy eigenvalues in eV (cm^{-1}) for the first five states of both wagging modes

Mode	$\omega_e(\text{cm}^{-1})$	$\omega_e x_e(\text{cm}^{-1})$	$\omega_e y_e(\text{cm}^{-1})$
in-plane	1263.27	12.9492	1.17117
out-of-plane	6.09491	-0.024239	-0.000590691

Table 6: parameters from fitting the exact vibrational spectrum to Eq. 3.2

Note that the out-of-plane motion is remarkably harmonic. Since at high vibrational excitation, the proton should spend more time near the classical turning points, the stretch motion will probably occur along an out-of-plane direction. With this in mind, we calculated the analogous potential curves at 50° out-of-plane. Except for a small shift of both the ground and excited states, the potential curves essentially overlay those taken along the O-H direction with no wagging motion.

We close this section with a brief discussion of calibration of all electron results with effective core potential results. We show in Fig 4 the potential curves obtained using the Stevens-Basch-Krauss effective core potential and the associated basis set, augmented with polarization functions (d functions on heavy atoms and p functions on the hydrogen atoms). We used the same level of CI in these calculations. We also show the all-electron results. For both sets of calculations, the

potential surfaces were shifted so the the $E(R_{eq.}) = 0$ for the ground electronic state. Note that, over the range calculated, This potential curve overlays the all electron potential surface, giving us confidence in using this ECP for large cluster calculations.

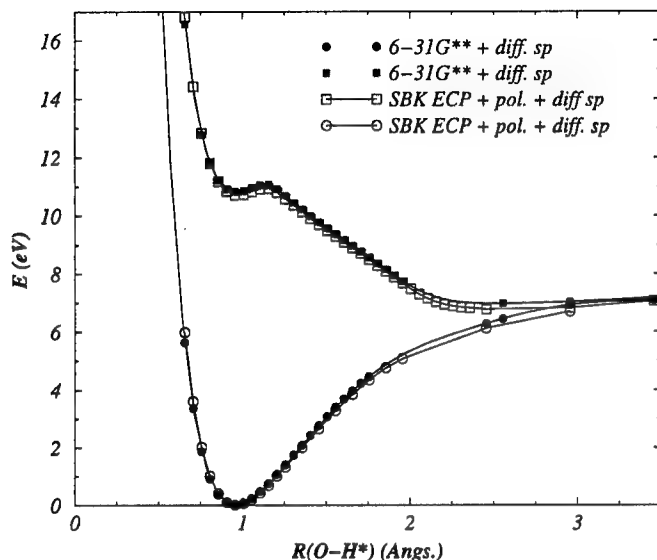


Figure 4: O-H Stretching mode potential surface for the all-electron and the SBK ECP calculations.

4 The Interface Proton

We see that at this level of approximation, on small clusters, the simple model discussed in Ref. [2] will not stand. We must stress, however, we may obtain different results using larger clusters. Previously, we have obtained an excitation energy of 8.9 eV on a $Si_5O_{16}H_{13}$ cluster at roughly the same level configuration interaction (All single and double excitations, [fill this in tomorrow]. Open shell Hartree Fock has proven more difficult here, but clearly that is the appropriate next calculation. During the current period, we have explored the possibility that the excitation is not from the silicon dioxide valence band, but, rather, from the silicon valence band. We have modeled this using the cluster shown in Fig. 5. Here we used the same all-electron basis set as for the bulk proton. We have obtained an equilibrium geometry within the Hartree-Fock approximation with and without the proton. The parameters are given in Table 7. The most striking result, evidenced in the equilibrium geometry, is that the positive charge prefers to reside in the oxide on the proton over residing in the silicon orbitals.

In Figs. 6 and 7 we present constant density surfaces for the HOMO and LUMO of the model

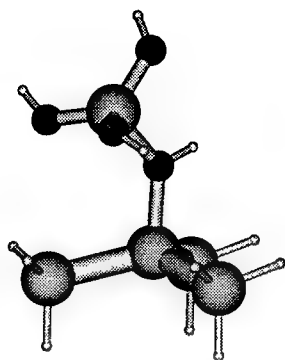


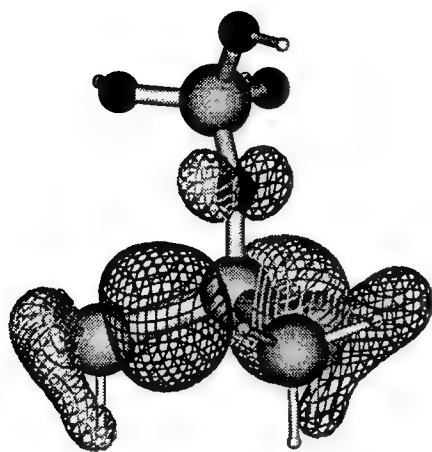
Figure 5: Cluster used for study of the interface proton.

interface without and with a proton, respectively.

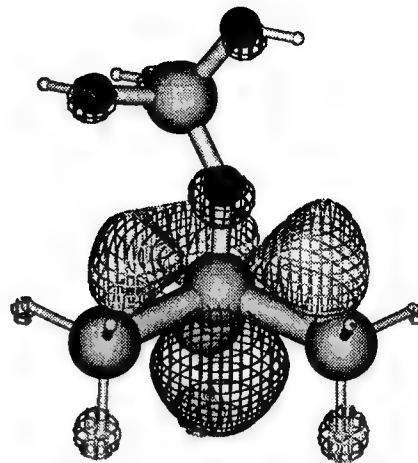
	<i>Si</i> – <i>SiO</i> ₂	<i>Si</i> – <i>SiO</i> ₂ with proton
$R_{Si_s-O*}(\text{\AA})$	1.66	1.90
$R_{Si_o-O*}(\text{\AA})$	1.62	1.75
$R_{Si_o-O_n}(\text{\AA})$	1.63	1.60
$R_{Si_s-Si_n}(\text{\AA})$	2.38	2.37
$R_{O*-H*}(\text{\AA})$		0.95
$\angle Si_o - O* - Si_s(^{\circ})$	148	134
$\angle O* - Si_o - O_n$	109.0	103.2
$\angle O* - Si_s - Si_n$	109.0	103.3
$\angle H* - O* - Si_s(^{\circ})$		115

Table 7: Equilibrium geometrical parameters for model interface with and without proton

For the model interface, both the HOMO and LUMO have large contributions from the silicon atoms simulating the bulk. This is consistent with known band gaps and band offsets for the Si-SiO₂ system. However, the band gap is far too large (12.5 eV). This overestimation, while typical of Hartree-Fock theory, is also due to fundamental quantum-confinement that is a direct result of the size of the cluster. On comparing Figs. 6 and 7, we see that, while the HOMO is mostly unperturbed by the presence of the proton, the LUMO is altered dramatically. It is dominated by proton-like



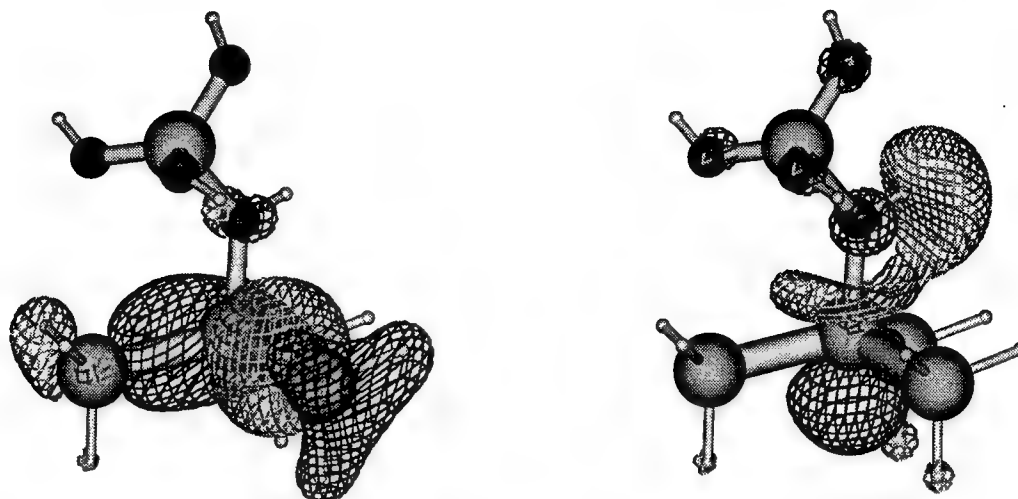
a. HOMO



b. LUMO

Figure 6: Constant density (0.04 \AA^{-3}) map for the frontier orbitals for the model interface.

density that has bonding interaction with the silicon surface orbitals. We are surprised that this is the LUMO, because in the bulk case, the proton excitation energy is essentially the SiO_2 band gap. We expect that in a larger cluster, where the silicon conduction band might be represented better, the proton state would be considerably above an LUMO that is similar to that shown in Fig 6 b. Using the same level of configuration interaction as we used for the bulk proton, we obtained an excitation energy of 6.46 eV. This reduction was expected because of the valence band offset between silicon and silicon dioxide. We have calculated the first excitation energy for the model interface to be 7.12 eV, compared to a 1.2 eV experimental value. As mentioned above, this large calculated value is in line with results on quantum confinement. So, this result can only be considered as preliminary. We are currently studying a much larger cluster. Never-the-less, the result leaves open the possibility that electrons are introduced into the SiO_2 from the silicon valence band through a variant of internal photoemission. Previously, this mechanism had been discounted because of the absence of observed compensation or annihilation of positive charge. In Sect. 5, we outline previous



a. HOMO

b. LUMO

Figure 7: Constant density (0.04 \AA^{-3}) map for the frontier orbitals for the interface proton.

arguments and how, using more realistic assumptions, we come to different conclusions.

5 Electrostatic calculations

In this section we outline the principle approximations and results from two sets of assumed boundary conditions and charge densities. In the past it has been argued that if the proton charge density were constant, so that the total positive charge were $2 \times 10^{12} \text{ cm}^{-2}$, and if the potentials at both boundaries of the oxide were $\phi(x) = 0$, then the field at either boundary of the oxide would be $\approx 5 \times 10^5 \text{ V/cm}$. This is trivially obtained by solving Poisson's equation in one dimension and is independent of the thickness of the oxide. However, the field is insufficient to determine whether electrons will remain trapped because the well depth is dependent on oxide thickness. For the case of a 1000 \AA oxide, the well depth would be 1.1 V and the electrons would be well trapped. Thus, even

if there were no trapping at protons, there would be compensation. One might conclude that, based on the observed absence of compensation, internal photoemission is not possible in this system. However, there are several problems with these assumptions. First, in the experiment of Vanheusden and co-workers, the proton distribution is not constant, but is assumed to be fairly compact, with all of the mobile protons very close to the substrate interface. Second, the assumed boundary conditions do not match the experiment. During exposure to UV-light, the oxide is uncovered so that the outer boundary would be considered to be free. If the sample sits on a metallic holder, the substrate is, to a reasonable approximation, grounded. However, there will be a depletion layer at the Si-SiO₂ interface that is not accounted for in the simplest model. This implies that there will be some potential drop across this depletion layer that will require calculation. We have considered the following system:

1. A 100 nm oxide with a constant proton density within 30 Å for the Si-SiO₂ interface so that the total density would be $1 \times 10^{12} \text{ cm}^{-2}$, with no proton density outside this region.
2. The outer SiO₂ boundary is assumed to have zero normal electric field.
3. The substrate is assumed to be grounded.
4. Substrate acceptor concentration of $1 \times 10^{16} \text{ cm}^{-3}$

Under these assumptions, we have solved Poisson's equation in the depletion approximation and compared it to results obtained with a one dimensional drift-diffusion device modeling code. The results are virtually indistinguishable, and are given in Fig. 8.

Several features bear notice. First, there is a large potential drop across the depletion region. Second, the proton distribution gives a very small potential barrier ($\approx 0.1 \text{ eV}$ for re-emission into the substrate. Finally, the free exterior boundary leads to no real potential well. Only the intrinsic oxide electron affinity ($\approx 0.9 \text{ eV}$) will prevent the electrons from exiting the top surface. It is likely, then, that most of the electrons that do enter the oxide conduction band will re-enter the silicon conduction band by surmounting the very small barrier induced by the proton density. From this calculation we conclude that photoemission into the SiO₂ does not necessarily lead to either proton annihilation, or compensation. Thus, the excitation channel posited in Sect. 4 appears plausible.

6 Discussion

The results we report here are both interesting and incomplete. The small cluster calculations on the bulk proton indicate that excitation from the silicon dioxide valence band requires much more energy

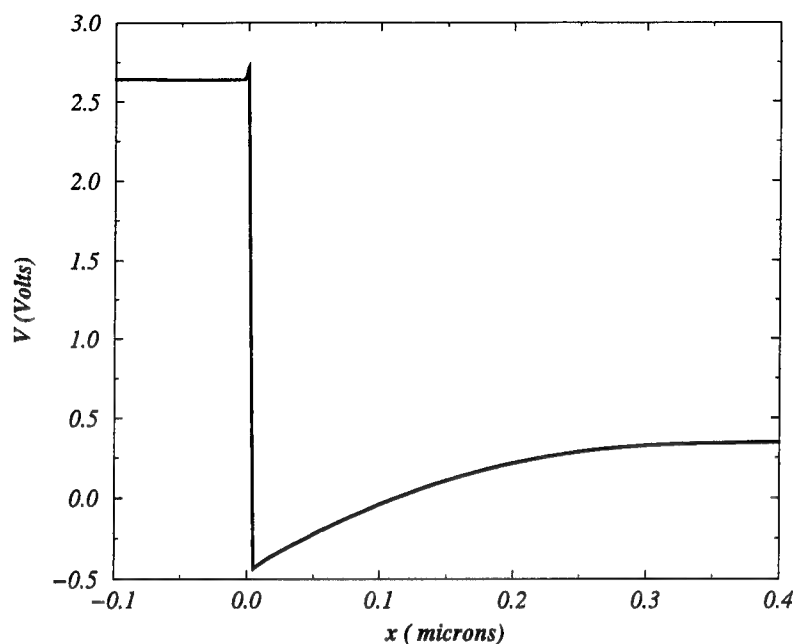


Figure 8: Electrostatic potential for $1 \times 10^{12} \text{cm}^{-2}$ protons within 30 Å of the Si-SiO₂ interface.

(*approx* 3 eV) than the experimental excitation. However, as we mentioned above, cluster size may play a crucial role. If, for the moment, we *do* assume that these results represent the solid faithfully, we are forced to conclude that the direct excitation model does not hold. With this possibility in mind, we report calculations on a small interface model that reflect some of the physics we might expect. The excitation energy is much lower and reflects the known offset between the silicon and silicon dioxide valence bands. Because of the poor representation of the silicon conduction band, and of the known quantum confinement in small silicon clusters, we will be required to use much larger clusters to obtain quantitatively meaningful excitation spectra. Of more immediate interest is the predicted stability of the proton. A positive charge would prefer to reside on the oxide side of the model interface in the form of a proton, than as a hole in the silicon valence band. We will address in the large cluster calculations whether an electron in the conduction band is predicted to destabilize the proton. If this holds, then the standard model for latent build-up of interface states must be revised. It is quite possible that protons are neutralized in irradiated oxides because of the presence of a large density of radiation-induced electron traps at the Si-SiO₂ interface.

7 Acknowledgments

I acknowledge with great pleasure the support of the group at Phillips Research site of AFRL, including Drs. Walter Shedd, Robert Pugh, B. K. Singaraju and S. P. Karna, during this summer fellowship. I also want to thank Dr. K. Van Heusden and Prof. G. Pacchioni for valuable conversations. GAMESS-UK is a package of ab initio programs written by M.F. Guest, J.H. van Lenthe, J. Kendrick, K. Schoffel and P. Sherwood, with contributions from R.D. Amos, R.J. Buenker, M. Dupuis, N.C. Handy, I.H. Hillier, P.J. Knowles, V. Bonacic-Koutecky, W. von Niessen, R.J. Harrison, A.P. Rendell, V.R. Saunders, and A.J. Stone. The package is derived from the original GAMESS code due to M. Dupuis, D. Spangler and J. Wendoloski, NRCC Software Catalog, Vol. 1, Program No. QG01 (GAMESS), 1980.

8 REFERENCES

1. K. Vanheusden, and A. Stesmans, *Appl. Phys. Lett.* 64, 2575 (1994).
2. K. Vanheusden, W. L. Warren, D. W. Fleetwood, S. P. Karna, A. M. Ferreira, R. D. Pugh, C. P. Brothers, B. K. Singaraju, R. A. B. Devine, and A. H. Edwards, *Appl. Phys. Lett.* 0, 0 (0).
3. K. Vanheusden, D. M. Fleetwood, J. R. Schwank, M. R. Shaneyfelt, T. L. Meisenheimer, and B. L. Draper, *IEEE Trans. Nuc. Sci.* NS- 45, XX (1998).
4. R. E. Stahlbush, and H. L. Hughes, in *35th Nuclear and Space Radiation Effects Conference*, edited by F. W. Sexton, (IEEE, Piscataway, 1998), p. 0.
5. A. H. Edwards, *Phys. Rev. Lett.* 71, 3190 (1993).
6. A. M. Ferreira, S. P. Karna, C. P. Brothers, R. D. Pugh, B. B. Singaraju, K. Vanheusden, W. L. Warren, and R. A. B. Devine, in *Amorphous and Crystalline Insulating Thin Films*, edited by W. L. Warren, R. A. B. Devine, M. Matsumura, S. Cristoloveanu, Y. Homma, and J. Kanicki, (Materials Research Society, Pittsburgh, 1997), p. 187.
7. S. P. Karna, R. D. Pugh, J. R. Chavez, W. M. Shedd, C. P. Brothers, B. K. Singaraju, M. Vitiello, and G. Pacchioni, *IEEE Trans. Nucl. Sci.* 45, XX (1998).
8. K. Vanheusden, W. L. Warren, and R. A. B. Devine, *J. Non-Cryst. Solids* 216, 116 (1997).
9. N. S. Saks, and D. B. Brown, *IEEE Trans. Nucl. Sci.* NS- 37, 1 (1990).
10. M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Natsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, *J. Comput. Chem.* 14, 1347 (1993).
11. W. J. Stevens, H. Basch, and M. Krauss, *J. Chem. Phys.* 81, 6026 (1984).

12. C. Moller, and M. S. Plesset, Phys. Rev. 46, 618 (1934).
13. K. Kano, Semiconductor Devices (Prentice Hall, Upper Saddle River, 1998)
14. K. Vanheusden, Private communication.

**INVESTIGATING THE USE OF OPTICAL FIBER AS
OPTICAL DELAY LINE FOR ADAPTIVE OPTICS SYSTEMS**

**Claudio Oliveira Egalon
Assistant Professor
Department of Physics**

**University of Puerto Rico
Mayaguez Campus
Mayaguez, PR 00681**

**Final Report for:
Summer Faculty Research Program
Phillips Research Site**

**Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC**

and

Phillips Research Site

September 1998

INVESTIGATING THE USE OF OPTICAL FIBER AS OPTICAL DELAY LINE FOR ADAPTIVE OPTICS SYSTEMS

Claudio Oliveira Egalon
Assistant Professor
Department of Physics
University of Puerto Rico

Abstract

Experiments were performed in order to determine whether optical fibers wound around piezoelectric tubes, PZT, which are connected in cascade, can be used as optical delay lines. The experiment performed consisted in winding two different lengths of fibers around two different PZT stretchers in cascade. The fibers were then spliced to each other and stretched. An optical interferometer was built in order to determine the resulting phase shift due to the stretching of the fibers. It was found, as expected, that the two stretchers added up to the number of fringes displaced in the interferometer. This simple experiment has demonstrated that multiple PZT stretchers may be used in cascade in order to achieve long Optical Phase Differences, OPD.

INVESTIGATING THE USE OF OPTICAL FIBER AS OPTICAL DELAY LINE FOR ADAPTIVE OPTICS SYSTEMS

Claudio Oliveira Egalon

Introduction

Optical fibers, as optical delay lines for adaptive optics, has been proposed by many in the scientific community. Its adoption could greatly simplify the adaptive optics systems that are currently in use. At present, adaptive optics systems use mirrors mounted in mechanical tracks as the sole means for creating an optical phase difference. These mountings are expensive, cumbersome and takes up considerable space. Optical fibers wrapped around PZT stretchers offer the possibility of replacing these complicated systems by creating as much optical phase difference, at a lower cost using a limited amount of space. In order to test the above assumption, experiments were performed using two optical fibers wrapped around two PZT stretchers. The fibers were connected in cascade and stretched. Their resulting phase shift was then measured using an optical interferometer configuration..

Experimental Procedure

A graphical description of the experimental set up is shown in Figure 1. A He-Ne laser ($\lambda=638$ nm) was used to excite two different fibers; a reference fiber and a fiber that was subjected to stretching by two different piezoelectric stretchers. A beam splitter was used to inject light in the two different fibers and, at the other end, a second beam splitter was used to recombine the beams from the fiber's output ends generating an interference

pattern, in the form of fringes, which could be observed with a camera. As the fiber in the stretchers was stressed, the fringes would move.

The first fiber had a total length of 3.13 m with 1.78 m wrapped around the PZT stretcher whereas the second fiber had a total length of 5.15 m with 1.78 m wrapped around the PZT stretcher number 2.

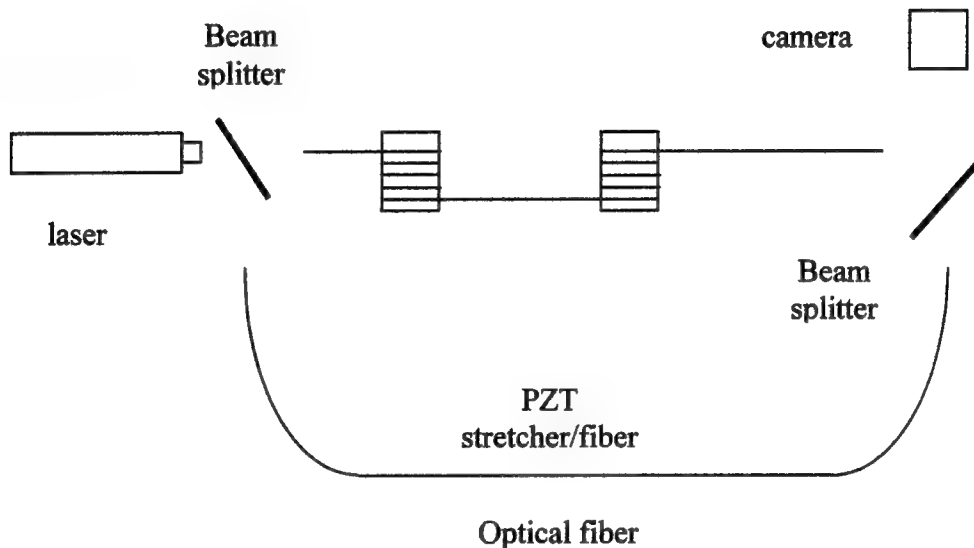


Figure 1. Experimental set up.

Experimental Results

The fiber stretchers were driven by a power supply (not shown in the Figure) which would incrementally stretch the fiber as voltage was slowly increased. Initially, a single PZT stretcher was connected to the power supply and the voltage varied from 0 to 180V resulting a total of 65 fringes displacements. The experiment was repeated several times in order to assure reproducibility and the number of fringes displaced for the number 1 stretcher remained within 65 and 70. The stretcher number 2 was then connected to the power supply resulting in 43 fringes displaced. Both stretchers were then connected in

parallel to the power supply resulting in the displacement of 100 fringes. The number of fringes displaced obtained whenever the two stretchers are connected agree with the sum of the number of fringe displacement whenever each stretcher is connected separately. For a laser wavelength of 638 nm 100 fringes would correspond to an optical phase difference of 63.8 microns or 0.0638 mm.

Conclusions

It has been demonstrated optical fibers wrapped around PZT's that are connected in cascade have their optical phase difference summed up. This result indicates that an optical fiber wrapped around several PZT's may generate even larger amounts of optical phase difference. In reality it is desired to obtain an OPD on the order of a few centimeters which could be achieved with a stretched length of fiber of roughly 2 km..

LOW LIGHT LEVEL ADAPTIVE OPTICS APPLIED TO VERY HIGH RESOLUTION IMAGING

Jefrey Foster Friedman
Associate Professor
Department of Physics

University of Puerto Rico Mayaguez
PO Box 9016
College Station
Mayaguez, PR 00681-9016

Final Report for
Summer Faculty Research Program
Phillips Research Site

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Phillips Research Site

August 1998

LOW LIGHT LEVEL ADAPTIVE OPTICS APPLIED TO VERY HIGH RESOLUTION IMAGING

Jeffrey Foster Friedman
Associate Professor
Department of Physics
University of Puerto Rico Mayaguez

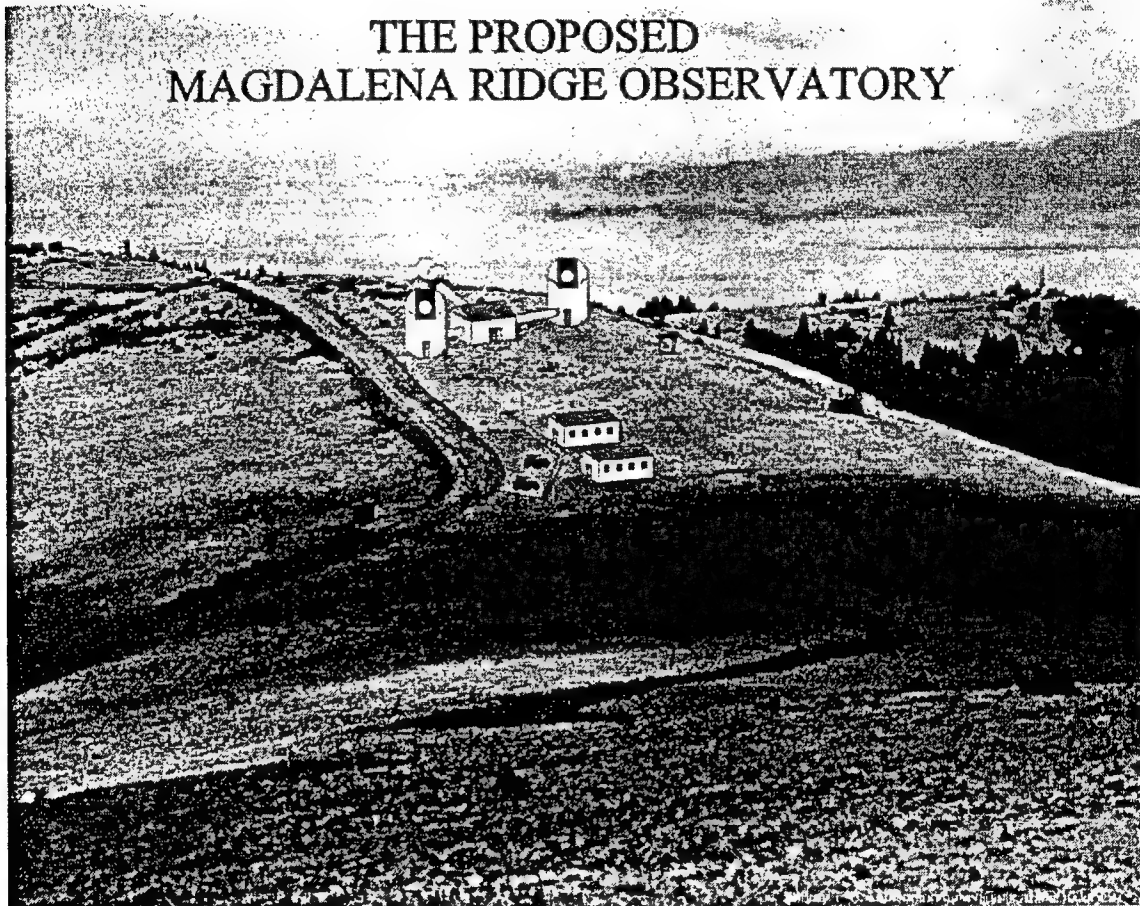
Abstract

Very High Resolution Optical Instruments (VHROI) typically have conjugate limited fields of view. To achieve this high resolution the light from dispersed elements must be accurately combined. The individual element beams must be in phase at the detector and have aligned and equivalent fields of view. Because of turbulence in the light path from target source to detector even the light from a moderately sized mirror is not coherent across the mirror. Separated elements introduce greater dislocations in the wavefront. If a high resolution image is to be created all of the dislocations in phase must be adequately compensated for. This generally requires an adaptive optics system. The shorter the wavelength, the larger the separation, and the larger the individual elements the greater the need for high order corrections by the adaptive optics. In this way a VHROI places very high constraints on the adaptive optics system. In addition there are other considerations that must be made in the physical design of the VHROI that have to do with adequate sampling of the source and minimizing the disturbances to the environment by the instrument and its infrastructure.

While our project is very concerned with all of these constraints this summers work has concentrated on the last two elements. The first issue is usually called U-V coverage of the dilute array, instantaneously and under earth rotation (for synthetic aperture construction) or active motion of the elements. Our system will most likely be static, but mobile elements are being designed for flexibility in the future. The second has to do with clean air over the elements. This can have several causes, of which we have only considered one in this study. This is the placing of the infrastructure for the individual elements in such a way as not to produce additional turbulence for elements down wind of the structures. This is accomplished by simply not placing structures along the same wind line. Other considerations in this class are turbulence of the structure on the enclosed telescope and turbulence generated by the telescope itself, either from its construction of heating/cooling properties. We have considered the temperature equilibrium characteristics of the primary mirrors for design purposes but have not considered them in this project.

Introduction

In the figure below the ridgeline is paralleled by the road and is a few degrees off of N-S. The prevailing winds are from the southwest which is the right in the figure. The Ridge is a photograph and the buildings are added to represent the proposed structures for the large telescopes and combining room of the MRO. The VLA is also to the right and is in the line of site from the ridge.



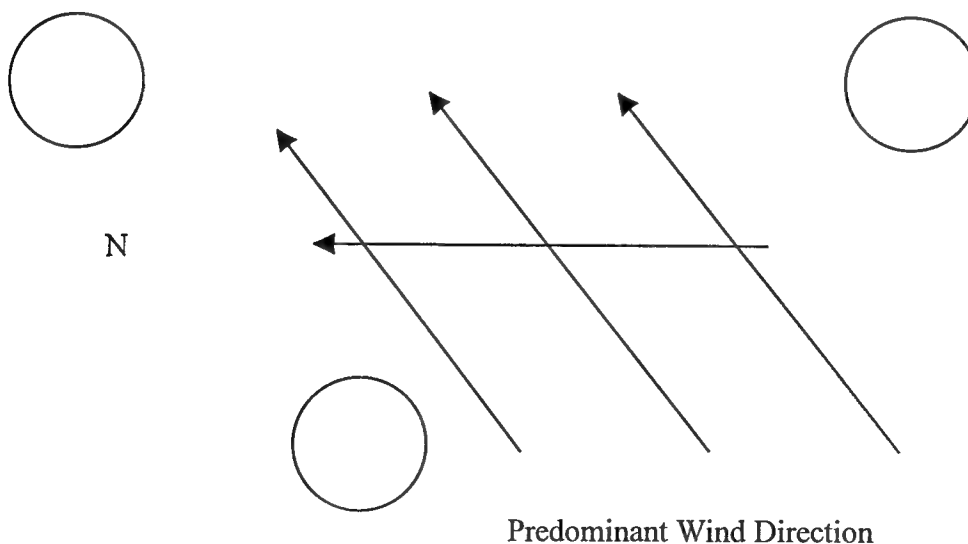
We are constructing the elements of a VHROI to be placed on this ridge in the near future. The University of Puerto Rico Mayaguez, with AFRL support is building three 0.75 meter telescopes on mobile robotic mounts with wavefront reconstructing Adaptive Optics (AO). The AO system has been developed at AFRL/DEBS over the past several years. The mirrors are of a fairly light-weight construction that assures rapid thermal

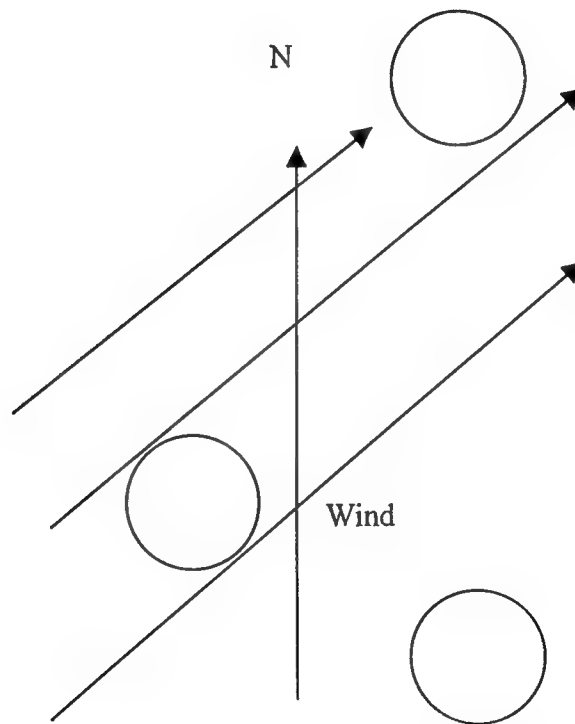
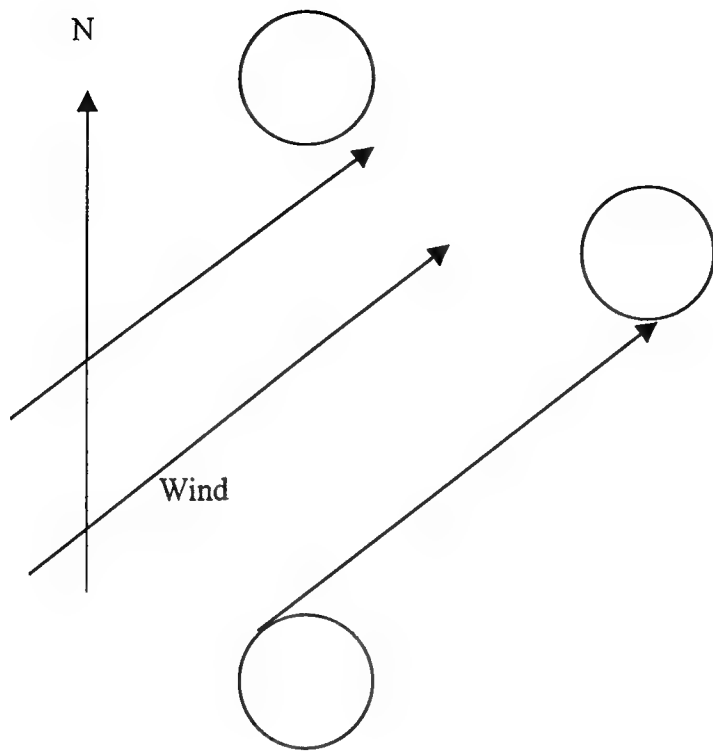
equilibration and are being built by Stabilite Optics. We are developing the mounts in Puerto Rico with Caribe Submarine and Space Technology using technology they developed for Satellite tracking stations. The mounts are robotic, mobile, remotely operable, self-aligning, self-positioning, and self-correcting. The combination of these three systems will yield a powerful mobile optical device capable of many surveillance needs. The combination of the three units will provide a versatile VHROI.

Plans for the Ridge include three 2.4 meter telescopes that will include AO and the capability to be combined into a VHROI as a group and in combination with the three 0.75 meter telescopes. It is the combination of 6 telescopes that we will study for u-v coverage and wind interference.

The space available on the ridge is very large and even kilometric baselines are possible. The problems of co-phasing optics 1000 meters apart are presently intractable but are open for future development. We have chosen a nominal 50 meters as the largest baseline for the 2.4 meter telescopes because it adds to the science that can be done from the Earth's surface and is reasonable to attempt co-phasing.

The following diagrams illustrate three nominal 30-40-50 meter arrays with the 50 meter baseline N-S to eliminate the longest delay line for combination. There is a penalty in u-v coverage, but it simplifies the combination and is very cost effective.





Infrastructure requirements dictate minimal telescope housing requirements. If we assume at least part-time use of the telescopes in the conventional astronomic or surveillance mode then they require complete freedom of movement and the minimum spacing, S, must be $2 \times D \times FN$; where D is the telescope primary's diameter and FN is the focal ratio (f number). In this case D is 2.4 meters, FN is close to 2. This gives $S=9.6$ meters and is an under estimate of the true requirement. We use a value of 12 meters for the large telescopes and 6 meters for the small ones. We use 9 meters for the small/large combination. This just says we have 12 meter domes for the large telescopes and 6 meter domes for the small ones. The minimum spacing is just that required abutting the domes. We have considered straight wind lines without turbulence or broadening for the first order approximation. It can be seen in the diagrams that the first situation is precluded and the next two are applicable with only slight modification. In the six telescope configuration first consideration is give to clean air over the large telescopes and then the smaller telescopes regardless of u-v coverage. This is because the usual mode of operation will be as single instruments and priority must be given to the largest mirrors. The Air Force interest in Very High Resolution (VHI) is secondary to simplicity, reliability, ease of use, and full AO compensation. The primary goal will be for single telescope surveillance and then VHR.

We will use studies performed by other groups before the final is determined. This has been a preliminary study to develop the tools and criteria necessary for a complete study. We will advance from Mathcad to IDL and develop a data analysis package once the configuration has been determined. It is likely the three small telescopes will remain as mobile units and enhance the capabilities and versatility of the VHROI.

I have included a program as written by Gary Loos of AFRL/DEBS with the input of Victor Gamiz, David Westpfahl, William Junor and myself. It is written in Mathcad for ease of use with a personal computer.

With the program we calculate the u-v coverage of the various arrays we produce using our selection criteria. Plots are produced that show the arrays of elements and the u-v coverage, instantaneously, with earth rotation, and as a function of 5 selected wavelengths. Taking data at several wavelengths effectively changes the baselines. Proper baseline units are wavelengths as the ratio of the baseline to observing wavelength determines the resolution.

I have included a single calculated example of the u-v coverage for illustration. We have performed the calculations for many configurations and will continue the effort through several hundred iterations. Also considered in the positioning of the individual telescopes are the possible light paths to the combining facility. This is necessary for construction considerations so as not to place a critical supporting member in the light path,

U-V COVERAGE OF A WAVELENGTH DIVERSE MRO ARRAY 14 SEPT. 1998

Baseline number of apertures T is (3 X 2.4 meter):

$$T := 3 \quad k := 0..T-1 \quad l := 0..T-1$$

Total number of apertures N is (plus 3 X 0.75 m):

$$N := 6 \quad i := 0..N-1 \quad j := 0..N-1$$

Array radius R (in meters) is: R := 20

Upper and Lower spectral band limits λ_{\max} and λ_{\min} (in meters) are:

$$\lambda_{\min} := 1 \cdot 10^{-6} \quad \lambda_{\max} := 3 \cdot 10^{-6} \quad \Delta\lambda := 2 \cdot 10^{-6}$$

Number of subspectral bands m is: 4 $m := 4$ $n := 0..m$

$$\lambda_0 := 1 \cdot 10^{-6} \quad \lambda_1 := 1.5 \cdot 10^{-6} \quad \lambda_2 := 2.0 \cdot 10^{-6} \quad \lambda_3 := 2.5 \cdot 10^{-6} \quad \lambda_4 := 3 \cdot 10^{-6}$$

$$\text{Latitude } \phi \text{ of 34 degrees is 0.593 radians} \quad \phi := 0.593$$

Primary telescope (3 X 2.4 meter) coordinates are:

$$y_0 := -15. \quad z_0 := 20. \quad y_1 := 15. \quad z_1 := 20. \quad y_2 := 15. \quad z_2 := -20.$$

Secondary telescope (3 X 0.75 meter) coordinates are:

$$y_3 := 0. \quad z_3 := 20. \quad y_4 := 0. \quad z_4 := 0. \quad y_5 := 15. \quad z_5 := 0.$$

Define positions Px,Py: $Py_i := y_i$ $Pz_i := z_i$ $Py'_k := y_k$ $Pz'_k := z_k$

Define baseline vectors X,Y in the ground plane.

$$X_{n \cdot N^2 + (j \cdot N + i)} := 0 \quad Y_{n \cdot N^2 + (j \cdot N + i)} := \frac{y_i - y_j}{\lambda_n} \quad Z_{n \cdot N^2 + (j \cdot N + i)} := \frac{z_i - z_j}{\lambda_n}$$

$$X'_{n \cdot T^2 + (l \cdot T + k)} := 0 \quad Y'_{n \cdot T^2 + (l \cdot T + k)} := \frac{y_k - y_l}{\lambda_n} \quad Z'_{n \cdot T^2 + (l \cdot T + k)} := \frac{z_k - z_l}{\lambda_n}$$

Transformation from ground coordiantes to celestial:

$$\begin{bmatrix} X'_{n \cdot T^2 + (1 \cdot T + k)} \\ Y'_{n \cdot T^2 + (1 \cdot T + k)} \\ Z'_{n \cdot T^2 + (1 \cdot T + k)} \end{bmatrix} := \begin{bmatrix} \cos(\phi) & 0 & -\sin(\phi) \\ 0 & 1 & 0 \\ \sin(\phi) & 0 & \cos(\phi) \end{bmatrix} \cdot \begin{bmatrix} X'_{n \cdot T^2 + (1 \cdot T + k)} \\ Y'_{n \cdot T^2 + (1 \cdot T + k)} \\ Z'_{n \cdot T^2 + (1 \cdot T + k)} \end{bmatrix}$$

$$\begin{bmatrix} X_{n \cdot N^2 + (j \cdot N + i)} \\ Y_{n \cdot N^2 + (j \cdot N + i)} \\ Z_{n \cdot N^2 + (j \cdot N + i)} \end{bmatrix} := \begin{bmatrix} \cos(\phi) & 0 & -\sin(\phi) \\ 0 & 1 & 0 \\ \sin(\phi) & 0 & \cos(\phi) \end{bmatrix} \cdot \begin{bmatrix} X_{n \cdot N^2 + (j \cdot N + i)} \\ Y_{n \cdot N^2 + (j \cdot N + i)} \\ Z_{n \cdot N^2 + (j \cdot N + i)} \end{bmatrix}$$

Define hour and declination angles θ and δ :

$$\text{pts} := 35 \quad h := 0.. \text{pts} \quad d := 0.. \text{pts} \quad \theta_0 := 0 \quad \delta_0 := \frac{\pi}{2}$$

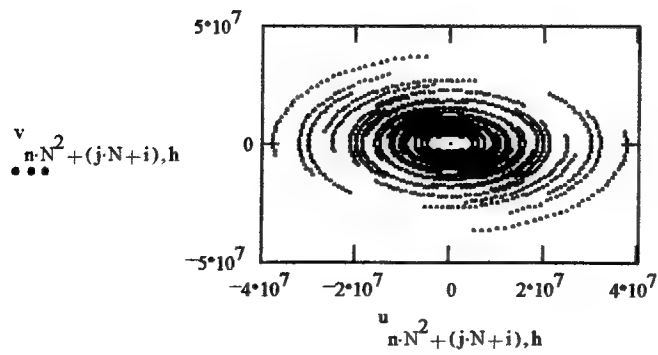
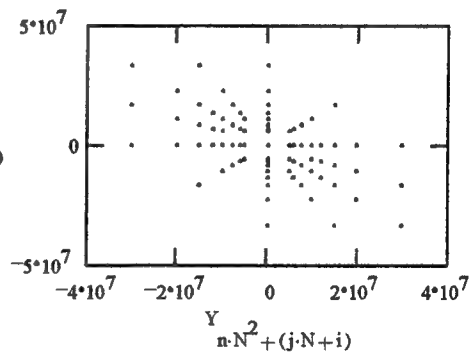
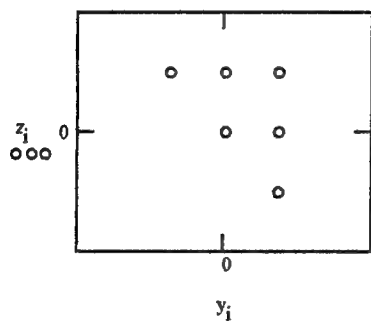
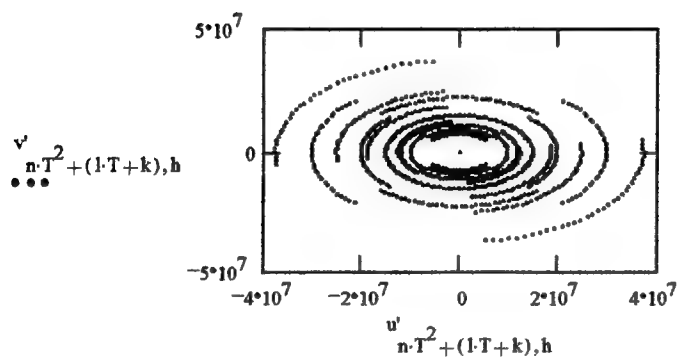
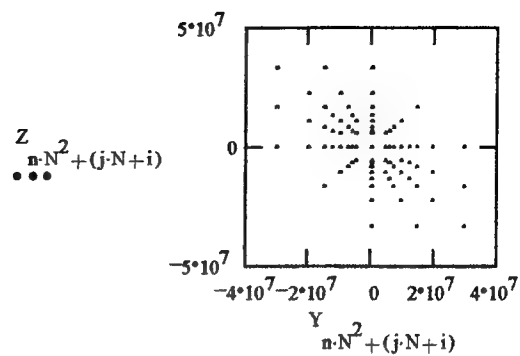
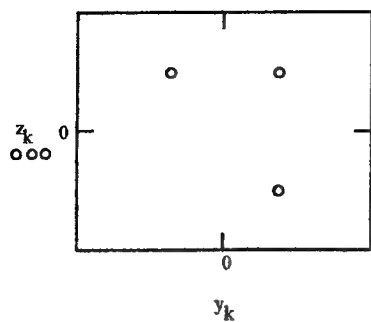
$$\theta_h := -\frac{\pi}{4} + h \cdot \frac{\pi}{70} \quad \theta_0 := -0.785 \quad \theta_{\text{pts}} := 0.785$$

The "Westpfahlian" transformation: (after David Westpfahl)

$$\begin{bmatrix} u'_{n \cdot T^2 + (1 \cdot T + k), h} \\ v'_{n \cdot T^2 + (1 \cdot T + k), h} \\ w'_{n \cdot T^2 + (1 \cdot T + k), h} \end{bmatrix} := \begin{bmatrix} \sin(\theta_h) & \cos(\theta_h) & 0 \\ -\sin(\delta_0) \cdot \cos(\theta_h) & \sin(\delta_0) \cdot \sin(\theta_h) & \cos(\delta_0) \\ \cos(\delta_0) \cdot \cos(\theta_h) & -\cos(\delta_0) \cdot \sin(\theta_h) & \sin(\delta_0) \end{bmatrix} \cdot \begin{bmatrix} X'_{n \cdot T^2 + (1 \cdot T + k)} \\ Y'_{n \cdot T^2 + (1 \cdot T + k)} \\ Z'_{n \cdot T^2 + (1 \cdot T + k)} \end{bmatrix}$$

$$\begin{bmatrix} u_{n \cdot N^2 + (j \cdot N + i), h} \\ v_{n \cdot N^2 + (j \cdot N + i), h} \\ w_{n \cdot N^2 + (j \cdot N + i), h} \end{bmatrix} := \begin{bmatrix} \sin(\theta_h) & \cos(\theta_h) & 0 \\ -\sin(\delta_0) \cdot \cos(\theta_h) & \sin(\delta_0) \cdot \sin(\theta_h) & \cos(\delta_0) \\ \cos(\delta_0) \cdot \cos(\theta_h) & -\cos(\delta_0) \cdot \sin(\theta_h) & \sin(\delta_0) \end{bmatrix} \cdot \begin{bmatrix} X_{n \cdot N^2 + (j \cdot N + i)} \\ Y_{n \cdot N^2 + (j \cdot N + i)} \\ Z_{n \cdot N^2 + (j \cdot N + i)} \end{bmatrix}$$

$$\text{wavelengths} := 5 \quad \text{small_array} := T \quad \text{large_array} := N$$



**ENVIRONMENTALLY-BENIGN SYNTHESIS OF
1,5-HEXADIYNE AND RELATED STUDIES**

**Vincent P. Giannamore
Associate Professor
Department of Chemistry**

**Xavier University of Louisiana
7325 Palmetto Street
New Orleans, LA 70125**

**Final Report for
Summer Faculty Research Program
Phillips Research Site**

**Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC**

and

Phillips Research Site

August, 1998

ENVIRONMENTALLY-BENIGN SYNTHESIS OF 1,5-HEXADIYNE AND RELATED STUDIES

Vincent P. Giannamore
Associate Professor
Department of Chemistry
Xavier University of Louisiana

Abstract

The compound 1,5-hexadiyne was synthesized from 1,5-hexadiene in two steps by a novel method that eliminated the use of molecular bromine, halogenated solvents, and liquid ammonia. All of these materials are used in the classic method. The initial step is the conversion of the hexadiene to 1,2,5,6-tetrabromohexane using sodium perborate and sodium bromide in glacial acetic acid instead of the standard method of treating the diene with bromine in a halogenated solvent. The final step was the phase-transfer catalyzed elimination of four equivalents of HBr from the tetrabromohexane using a base derived from potassium hydroxide and pinacol, instead of the usual sodium amide in liquid ammonia. The synthesis proved successful in producing 1,5-hexadiyne. However, more work needs to be done to increase the yield and purity of the final product. In addition to the synthetic work, calculations comparing the materials costs for the first step with its traditional counterpart were performed. Finally, theoretical computations of the enthalpy of formation for a number of compounds of interest were performed.

ENVIRONMENTALLY-BENIGN SYNTHESIS OF 1,5-HEXADIYNE AND RELATED STUDIES

Vincent P. Giannamore

Introduction

Interest in the use of acetylenic hydrocarbons as rocket fuels has a long history. {1, 2} Recently, because of renewed interest in 1,5-hexadiyne, a need for large quantities of this compound has developed. However, the long-time supplier of this compound has been acquired by another company and no longer produces it. Therefore, the Air Force Research Laboratory has undertaken to produce the material in quantity. The traditional laboratory method for the production of this compound is a two-step process, which involves the conversion of the readily available 1,5-hexadiene to 1,2,5,6-tetrabromohexane and the subsequent transformation of that compound into 1,5-hexadiyne. {3, 4, 5}

The first step of the traditional method involves the use of molecular bromine, which is known to be deleterious to the environment as well as harmful to human health in other ways. While the reaction can be performed in diethyl ether, it works better in a halogenated solvent such as dichloromethane. Halogenated solvents, of course, present environmental and health problems also. In addition, dichloromethane is attacked by strong bases. Since strong bases are used in the next step, the solvent must be completely removed before then.

After production of 1,2,5,6-tetrabromohexane in step 1, it is converted to 1,5-hexadiyne by a double double dehydrohalogenation reaction in which four molecules of HBr are eliminated from each molecule of the substrate. This elimination requires a very strong base and has traditionally been done using an alkali amide in liquid ammonia. Once again, both environmental and health risks are associated with the procedure.

All of the risks associated with the above processes are magnified when working on a larger scale. Materials such as bromine and liquid ammonia also require special handling, apart from those necessitated by environmental considerations. For these reasons, we sought to develop a synthesis of 1,5-hexadiyne from 1,5-hexadiene that was more environmentally benign and easier on the operator than the traditional method. Work on this synthesis is the

major subject of this report. We also report on relative costs of materials for part of the synthesis. Finally, theoretical computations of the enthalpy of formation for some related acetylenic and allenic compounds are reported.

Discussion of the Problem

A recent communication by Kabalka, Yang, Reddy and Narayana {6} reported a new method for bromination of alkenes that does not involve the use of either molecular bromine or halogenated hydrocarbons. Instead, these workers used sodium perborate and sodium bromide in glacial acetic acid to accomplish the bromination. Sodium perborate, a mild oxidizing agent, oxidizes the bromide in the presence of the alkene resulting in addition of bromine to the double bond. We sought to apply this method to the bromination of 1,5-hexadiene. The problems to be solved at this stage can be divided into three parts. First, since the synthesis had only been reported in a communication, experimental details were sparse. None of the brominations reported in the paper were done on a diene. So, many of the basic experimental details had to be worked out. Second, work-up of the product as reported in the paper involved the use of column chromatography in the final purification step. This would be impractical upon scale up, so a different method had to be found. The third part is related to the second. Kabalka and coworkers worked on a small scale and used quantities of glacial acetic acid on the order of 25 mL. On the scale on which we were working, much more acid was being used and, upon scale up, several liters might be required. Besides problems in separating the product from this solvent on that scale, it also seemed desirable to find an alternative solvent for larger scale work to avoid working with large quantities of a concentrated acid. It should also be noted that, although we have been referring to 1,2,5,6-tetrabromohexane as a single compound, it actually exists as three stereoisomers: one meso compound and a pair of enantiomers that are diastereomeric to the meso compound. Since all of these stereoisomers should be convertible to the diyne by the same procedure, the existence of several stereoisomers does not actually present a problem and separation of the stereoisomers from each other was not required.

Dehmlow and coworkers {7, 8} have developed a method for the dehydrohalogenation of 1,2-dihalides to 1-alkynes via phase transfer catalysis. This procedure avoids the use of both alkali amides and liquid ammonia. Their method

involves the conversion of a bis-tertiary diol such as pinacol (2,5-dimethyl-2,5-hexanediol) into its monoanion salt through reaction with powdered potassium hydroxide. The monoanion is a strong base but a weak nucleophile. This makes it an ideal reagent for elimination reactions. The anion is carried into the organic phase by the phase transfer catalyst. It acts as the base in the dehydrohalogenation, regenerating pinacol in the process. Thus, the pinacol anion is the effective base, but it is needed in only catalytic amounts. It is the potassium hydroxide that is used up in the reaction. The elimination takes place in a hydrocarbon solvent. Thus, the use of both alkali amides and liquid ammonia is avoided.

Several problems needed to be solved in order to apply this reaction to the synthesis in hand. One problem is the time and temperature needed for the procedure. Dehmlow, Thieser, Sasson and Neumann {8} reported yields in the 80% to 90% range with reaction times of less than an hour for most alkynes that they synthesized. Phenylacetylene was an exception, requiring 4.5 hours, as was acetylene itself, which required five hours and was formed in only 60% yield. Most eliminations were done at 80-90 °C, with none over 110 °C. On the other hand, Brandsma {3} reports that higher temperatures are required for smoother dehydrohalogenations in some cases. At lower temperatures, the substrate is only incompletely converted to alkyne. He generally uses slightly longer times than Dehmlow. None of the reported reactions produced a diyne. It seemed reasonable to assume that synthesis of a diyne would require some combination of a longer time and a higher temperature.

Another problem to be solved was the optimal amount of KOH for the reaction. Dehmlow used a 10:1 mole ratio of KOH to substrate; Brandsma used a 12:1 ratio in one case and 18:1 in the another case. The larger ratio was used in a synthesis that also required a higher temperature to complete the dehydrohalogenation. In the case of diynes, therefore, ratios on the order of 20:1 to as high as 36:1 would be called for if analogous conditions were used. Even on the laboratory scale, but especially for scale up, it seemed desirable to cut down on the amount of potassium hydroxide, if possible. This is especially true since the reaction gives optimum results only when freshly machine-powdered KOH is used.

For volatile alkynes, such as 1,5-hexadiyne (b.p. = 86 °C), both Dehmlow and Brandsma state that the product can be distilled directly from the reaction mixture when a high boiling solvent is employed. Therefore, it would seem that work up should be simple. The 1,5-hexadiyne should distill from the reaction mixture as it is formed. When that distillation stops, the reaction is complete. The distillate is then redistilled to give the final purified product. In practice, this proved to be much more complicated. Much work had to be done to develop a work up procedure for the reaction.

While the new synthesis method would have advantages if it could be perfected, the cost of the method in comparison with the traditional method must be considered. There are many factors involved here, including waste disposal, cost of complying with legal requirements, and others. As a first step, we calculated the cost of materials for the preparation of 1,2,5,6-tetrabromohexane by both routes. The results are reported below.

Finally, there are a number of acetylenic and related compounds that are potential synthetic targets by methods being developed. The selection of actual targets for synthesis must be guided by practical considerations. One of these is the energetics of the compound. A convenient way of estimating this is by the groupadditivity method first developed by Benson and Buss {9}. Using this method, as modified by more recent work {10, 11}, we calculated the enthalpy of formation for some potential targets. These results are reported.

Methodology

Our approach to the synthetic problem was to work on both steps concurrently. Thus 1,2,5,6-tetrabromohexane was prepared by both the classical method using molecular bromine and by the new method. By doing this we were able to supply sufficient starting material for the second step of the synthesis. In this manner, work on the second step did not have to be held up while problems with the first step were being worked on. This also allowed us to make a practical comparison between the two methods for step one.

The synthesis of the tetrabromide by Kabalka's method was performed six times with various modifications each time as reported below. We began with the procedure followed by Kabalka and coworkers for

1,2-dibromocyclohexane. They added 30.0 mmol of NaBr to a mixture of 15.0 mmol of $\text{NaBO}_3 \cdot 4\text{H}_2\text{O}$ and 13.5 mmol of cyclohexene in 25 mL glacial acetic acid and stirred for two hours. Then, they diluted the mixture with water, extracted the product with ether, and dried it over MgSO_4 . The solvent was removed by column chromatography to give the product in 87% yield.

In our first experimental run, we began by increasing the mole ratios of the inorganic reagents to the alkene. Kabalka used 2.22 moles of sodium bromide and 1.11 moles of sodium perborate per mole of double bond. We increased the ratios to 2.5:1 and 1.5:1, respectively. These are amounts per double bond. Since we were working with a diene, the ratios of moles of reagents to moles of substrate were actually twice as large. At the same time, we decreased the relative amount of acetic acid. One way to measure this is to consider the volume of solvent per mass of other starting materials. Kabalka and coworkers used 25 mL of acetic acid per 6.46 g of material. On the scale on which we were operating in this experiment, 129.5 mL of acetic acid would have been a comparable amount. We attempted to run the reaction using 100 mL. The reaction mixture is heterogeneous from the beginning. As time went on, more and more solids were formed and the mixture thickened and could not be stirred. It became necessary to add more acetic acid to keep the mixture stirring. TLC showed that the reaction was not yet complete at this point, although two hours had elapsed. In this run, the reaction was being performed under a nitrogen blanket. However, the thickening of the mixture to the extent that it impeded stirring was later observed even without the nitrogen flow. It appears that the ratio of 25 mL/6.46g gives approximately the minimum amount of acetic acid required. The yield in this run was 85%, which was not an improvement over Kabalka's results. Because of this, combined with the fact that using higher ratios of reagents to diene would require more acetic acid, we used Kabalka's ratios in subsequent runs. The work up will be discussed separately below.

In the second run of the reaction, most of the changes were in the work up, which will be discussed below. The initial amount of acetic acid was increased slightly, although the total amount was less since no additional acid was required as it was in run 1. The reaction continued to stir over the weekend. The reaction was successful with the ratios of reagents to double bonds equal to those in Kabalka's paper. Although the yield was lower, it was believed that this was due to the work up.

Experimental runs 3 and 4 used amount of reagents and solvent calculated according to the ratios used by Kabalka. Nitrogen flow during the reaction had been discontinued after run 1, but reagents had been added under a nitrogen blanket and the system closed beginning with run 2. By run 5, nitrogen was not used at all. Although sodium bromide is hygroscopic, it does not appear critical, at least in the desert, to maintain a dry atmosphere as long as time is not wasted in starting the reaction.

Experimental run 5 was done on a 2.5 times larger scale than the previous runs. An attempt was made to cut down slightly on the relative amount of acetic acid, but once again the mixture thickened and more acid had to be added. In run 1, the reaction had been noted to be very slightly exothermic for a short time at a point before the stirring had stopped. This had not been noted in the subsequent runs, but in run 5, on the larger scale, it was once again noted. This time the heat was much greater, so an ice bath was applied. After a short period, the flask cooled and did not become hot again. The initial mixture is white. After stirring for a while, it becomes yellow. It was at about the same time that the exothermicity was noted in both cases. Other than the flask feeling hot, no other evidence of exothermicity was noted. For example, there was no boiling noted, and the top was not blown off of the closed flask.

Experimental run 6 was on a scale three times greater than run 5. It was performed in a reaction kettle with mechanical stirring, instead of in a flask with magnetic stirring as had been done previously. An ice bath was applied from the beginning this time and the kettle was fitted with a reflux condenser. No exothermicity was noticed under these conditions. However, the point at which the color change from white to yellow occurred came overnight. Still, there was no evidence in the upper regions of the kettle or in the condenser of any reflux or other results of an exothermic process. The ice bath had been allowed to melt and warm up overnight. The mixture stirred for some time overnight and then over the weekend at room temperature with no ill effects. It appears that the exothermicity may be occurring only momentarily during the reaction.

It seems reasonable to assume that the yellow color is due to excess bromine. As the perborate oxidizes the bromide to bromine, it adds to the double bond. Since the perborate and bromide are slightly in excess, the total amount of bromine formed is in excess of the amount that can be added to the double bonds. The yellow color may be due to

this excess, or it may be that production of bromine by the oxidation outpaces the consumption of bromine by the addition reaction so that the yellow color first appears when a temporary excess of bromine is present. Both reactions are relatively fast processes, occurring almost instantaneously at room temperature. It is not known which reaction is faster. The color never gets any darker than a bright yellow. So, a large excess of bromine is never present.

When sodium perborate and sodium bromide are stirred in glacial acetic acid in the absence of an alkene, the yellow color begins to form immediately and steadily darkens to the typical reddish brown bromine color. Since the yellow color does not appear for some time in the presence of the diene, it may be that the addition is faster than the oxidation and the yellow color only appears when the bromine formed in the oxidation step no longer has any double bonds to add to. In that case, the persistence of the yellow color would be a marker for the completion of the reaction. However, when molecular bromine was added to 1,5-hexadiene, the bromine color appeared to persist when TLC indicated that the reaction had not gone to completion yet. TLC also indicated that tetrabromide formation was not yet complete when the yellow color appeared in the newer method. It appears that the reaction takes longer than the two hours that the simple alkenes require according to Kabalka's procedure. In all cases, we ended up stirring the mixture overnight and the reaction usually appeared complete at that point as judged by TLC. It may be worthwhile to follow the progress of the reaction much more closely by TLC or even by quenching and working up a series of small-scale reactions at regular intervals. This could either save wasted time which results in little or no increase in yield or indicate that additional reaction time is warranted by a significant increase in yield.

Some attempt was made to discover if another solvent might be suitable for the reaction. In the addition of bromine to alkenes using molecular bromine, potentially nucleophilic solvents, such as water or alcohols, must be avoided because the solvent usurps the place of the second bromine atom in the addition forming halohydrins (in the case of water) instead of dibromides. We chose to test dichloromethane and dimethyl sulfoxide as representative nonpolar and polar aprotic solvents respectively. These solvents did not dissolve a mixture of sodium bromide and sodium perborate that readily dissolved in the same amount of glacial acetic acid. We also tried adding glacial acetic acid to the other solvents, but the reagents still did not dissolve and no evidence of bromine formation was noted.

The first step in the work up used by Kabalka and coworkers was to dilute the reaction mixture with water. In our case we accomplished this by pouring the mixture into an equal volume of water. The mixture contained a large quantity of undissolved solid. This solid is soluble in ether and most, if not all, of it turned out to be product. The product is soluble in acetic acid to some extent, however. So, there is still some product in solution that must be recovered.

Initially, the diluted reaction mixture was extracted with ether. The layers did not separate well, however, even upon standing for a long time. It would also take several ether extractions before all of the solid dissolved and the undissolved solid made the interface between layers difficult to see. When the solid is recovered by filtering the reaction mixture before the ether extraction, less ether is required to dissolve the solid and it does not complicate the work up of the liquid portion. The initial portion of ether should be relatively large. For example, in extracting 600 mL of diluted reaction mixture with 300 mL of ether, it was planned to use three 100-mL portions. However, the first portion did not separate even upon standing. When the extractions were done with a 200-mL portion followed by a 100-mL portion, there were no problems of this nature.

Removal of the acetic acid was the next problem. In the example cited above, the volume of the organic layers after extracting with 300 mL of ether was approximately 400 mL. While some increase in volume might be attributable to extracted product and to the partial solubility of water in ether, it seems probable that most of this is acetic acid. Even the solid portion, however, still had an acetic acid odor, after several water washings. Washing with aqueous sodium bicarbonate was tried at first. However, with the volume of acetic acid present, many washes were required before gas evolution ceased and the washes were basic to pH paper. Even after all of this, the solid product at the end usually had an acetic acid odor. The bicarbonate washes were followed by washes with water and saturated aqueous sodium chloride solution. After drying over MgSO_4 , the solvent was stripped on a rotary evaporator. Another procedure used to remove acetic acid omitted any wash with bicarbonate solution. The ether extract was washed with water and saturated sodium chloride only. When the ether extract was washed with water, the volume of the organic layer decreased. In the example cited above, the 400-mL ether solution was reduced to about 300 mL. Of course, some ether is soluble in water, but it appears that much of the acetic acid is removed in this wash. The

product is not soluble in water. Still, after the ether is ultimately stripped on the rotary evaporator, after drying, acetic acid odor is still present. The following procedure was helpful, but still not completely effective. The ether was stripped on the rotary evaporator. After no more ether was present, the temperature of the water bath was raised and as much acetic acid as possible was removed on the rotary evaporator. Then, the flask was transferred to a high vacuum line at a nominal pressure of 20 to 25 millitorrs to remove remaining acetic acid. This was not completely successful either. We were able to obtain a dry, white solid product, but the vinegary odor was still present. Probably a combination of the bicarbonate washes and the rotary evaporator-high vacuum line procedure would be most effective. The final recommended procedure for the reaction and work up is given in the results section below.

The preparation of 1,5-hexadiyne from 1,2,5,6-tetrabromohexane by Dehmloew's method was attempted eight times. While some product was obtained each time as evidenced by either the boiling point or spectral evidence, both the yield and the purity of the product must be improved. It appears that part of the problem in both cases is that the tetrabromide is not completely converted to diyne. Both IR and NMR evidence indicate that a vinyl bromide is present along with terminal alkyne. Gas chromatography shows the samples to be mixtures. However, no matter which component of the mixture is the major one, there are still strong peaks in the IR indicating terminal alkynes and weaker ones indicating a double bond. The NMR also shows weak signals for vinyl protons and much stronger signals attributable to the diyne. This could be explained if some of the tetrabromide is being completely converted to the diyne, while a smaller portion is undergoing a double dehydrohalogenation on one side, but only a single elimination on the other side. This would produce $\text{HC}\equiv\text{CCH}_2\text{CH}_2\text{CH}=\text{CHBr}$ and/or $\text{HC}\equiv\text{CCH}_2\text{CH}_2\text{CBr}=\text{CH}_2$ as well as $\text{HC}\equiv\text{CCH}_2\text{CH}_2\text{C}\equiv\text{CH}$. Thus, it appears that one or both vinyl bromides are being produced along with the alkyne and not being completely separated from it in the work up.

The problem of incomplete reaction might be attributed to one of three causes. One of these is the temperature. Brandsma notes that some dibromides are not smoothly dehydrohalogenated except at higher temperatures and states that incomplete reaction is the cause. Related to this is the time factor. The reaction must be allowed to proceed for a long enough time to complete the eliminations, regardless of the temperature. Finally, the ratios of the reagents

to one another must be satisfactory. In this procedure, pinacol and tetraoctylammonium bromide are needed in catalytic amounts only. It may be of some help to increase the amounts to be certain that enough of each is available in the reaction mixture. However, the principal concern is the ratio of potassium hydroxide to tetrabromide.

Initially, we set up the apparatus so that the relatively volatile 1,5-hexadiyne (b.p. = 86 °C) would distill from the reaction mixture as it was formed. To this end, we used a high boiling hydrocarbon, dodecane, as solvent. Although dodecane boils at 216 °C, there was usually GC and spectral evidence that some of it was in the product mixture regardless of the work up method. Switching to Isopar V, a hydrocarbon mixture from Exxon Chemical that boils at 272 – 311 °C, reduced the amount of solvent contaminating the product to only trace amounts. However, the major problem with this method is that the incompletely reacted product was also distilling from the reaction mixture. Therefore, distilling the product directly from the reaction mixture as it is formed is not practical in this case. We also tried letting the reaction mixture reflux for awhile before distilling the crude product away. This did not result in pure diyne either. The time allowed to reflux may not have been long enough. However, the ratio of KOH to diene was also lower. It may be that lowering the amount of KOH counteracted the effect of allowing a longer reaction time. Results at least as good as those obtained by heating the mixture (either refluxing or distilling the product as formed) were obtained by allowing the reaction mixture to stir at room temperature. Again, this was only done with low KOH to diene ratios (5:1 or 8:1). It may be that this method would work well with larger amounts of KOH.

The order of addition of the various reagents was varied. Dehmloew added dibromide to a mixture of all of the other materials. After an initial exothermic reaction occurred, he then heated the mixture. Brandsma heated the KOH and solvent, then added the catalyst and the pinacol, and finally added the dibromide. We tried several variations, including adding the KOH last. There did not seem to be a reason to prefer one method to another as long as the stirring was efficient. However, adding the KOH first does keep it in excess at all times. If the amount of KOH proves critical, then it is probably advisable to add the tetrabromide to it rather than the reverse.

Distilling 1,5-hexadiyne appears to be a problem. One group of previous workers {1} reported adding ferrous sulfate and *p-tert*-butylcatechol to the compound before distilling it under nitrogen. At first, we attempted to distill it from the reaction mixture and then redistill it using a short-path low-hold-up still. In both of those distillations, the temperature would not remain steady at the boiling point. Instead, it would rise and fall over wide ranges. Ferrous sulfate and *p-tert*-butylcatechol were not immediately available to us.

To avoid any problems caused by heating the compound and keeping it at a temperature near its boiling point for a long period, we began distilling it on the high vacuum line. After the reaction was complete, the reaction mixture was filtered to remove unreacted KOH and other undissolved solids. It was then dried over MgSO_4 and then placed on the high vacuum line at nominal pressures of 20 to 25 millitorrs for several hours. We tried doing this distillation at elevated temperatures (35 °C and 55 °C), at room temperature, and at reduced temperature (-5 °C). In some cases, we took the distillate that was recovered from the liquid nitrogen cooled trap in this distillation and redistilled it on the same line at lower temperatures using a liquid nitrogen/acetone bath. This was done partially because some solvent was contaminating the first distillate. Unfortunately, dodecane has a lower freezing point than 1,5-hexadiyne, even though its boiling point is considerably higher. However, when we replaced dodecane with Isopar V as the solvent, the problem of solvent contamination was considerably decreased.

In one experiment, we distilled the same reaction mixture consecutively at -5 °C, room temperature, 35 °C, and 55 °C. The temperature was maintained at the given temperature for several hours and the distillate recovered from the trap before heating to the next higher temperature. Gas chromatography showed the first three distillates to be mixtures with three or four major components each. The solvent, Isopar V, was not one of these components. It was present only in trace quantities. No GC was obtained on the fourth distillate. The infrared spectra of all four distillates showed absorptions due to $\text{C}\equiv\text{C}$, $\equiv\text{C-H}$, and $\text{C}=\text{C}$. As the distillation temperature increased, the relative strength of these peaks compared to the aliphatic CH stretches decreased. Two points to be noted here are that the $\text{C}=\text{C}$ peak was present even at the lowest temperature and that there was still a $\equiv\text{C-H}$ peak in the distillate obtained at

the highest temperature. The final recommended procedure for the reaction and work up is given in the results section below.

For the cost comparison, we calculated the amount of material needed to produce 1 g of 1,2,5,6-tetrabromohexane using mole ratios employed in published procedures for each method. For solvents, we determined the amount based on the mass of the other materials using the same solvent volume to material mass ratio as the published procedure. We obtained the prices for each material from a recent Aldrich Chemical Company catalog using the most economical price listed for the grade of reagent used. The comparison is shown in the results section below.

Calculation of the enthalpy of formation for each of the compounds of interest was done using Benson's group additivity method. {9, 10} The group additivity values used came from those published by Cohen {11}. These are based on the latest experimental data, so that theoretical calculations are not based on theoretical data as has occasionally been done in the past. Essentially use of the method involves hypothetically breaking the molecule down into fragments. The enthalpy is calculated by adding the contribution from each fragment (the group additivity value) plus corrections for certain structural features. Thanks to Jeffrey Mills for advice on applying this method. The enthalpies of formation are presented in the results section.

Results and Conclusion

1. Synthesis of 1,2,5,6-tetrabromohexane. The following procedure is recommended for the preparation of this compound.

Quantities of materials should be determined as follows. The amount of 1,5-hexadiene used should be equimolar with the desired theoretical yield of the tetrabromohexane. In calculating the mass, remember that the diene is 97%. The mole ratio of sodium bromide to 1,5-hexadiene should be 60/13.5. For sodium perborate, the mole ratio should be 30/13.5. The amount of glacial acetic acid should be calculated by adding the masses of diene, sodium bromide and sodium perborate to be used and multiplying the sum by 25 mL/6.46 g. The mixture became impossible to stir magnetically any time less acetic acid was used. It may be that efficient mechanical

stirring may allow less acid to be used, but we suggest not decreasing the volume even with mechanical stirring.

To a reaction kettle or flask fitted with an efficient stirrer and a reflux condenser, add the 1,5-hexadiene, the sodium perborate, and approximately half of the glacial acetic acid. Stir the mixture at 175 rpm and add the sodium bromide followed by the remaining acetic acid. Increase the stirrer speed to 300 rpm and apply an ice-water bath. The mixture should be white at first, but eventually becomes yellow. Heat is evolved at about the same time as the color change occurs. The bath can be removed after the exotherm has subsided. The reflux condenser is probably not necessary after that point either. Continue stirring the mixture overnight and/or monitor the progress of the reaction by TLC.

Pour the reaction mixture into an equal volume of water. Filter the resulting mixture and save both the solid and liquid portions. Dissolve the solid in the minimum amount of diethyl ether. Wash it with water and then with saturated aqueous sodium chloride solution. Dry the solution over anhydrous magnesium sulfate. Strip the solvent on a rotary evaporator. This should result in a white solid. Allow this to air dry. If the odor of acetic acid is still evident, it may be placed on the high vacuum line as described below.

Take the filtrate from above and extract the product from it with diethyl ether. One may be tempted to combine this solution with the ether solution above. However, since the product from the solid is usually much better quality than the one obtained from the filtrate, this is not recommended. Wash the ether solution with aqueous sodium bicarbonate (10% or saturated) until the wash liquid is no longer acidic. Then, wash the solution with water followed by saturated aqueous sodium chloride solution. Dry it over anhydrous magnesium sulfate and strip the ether on a rotary evaporator. After the ether has been evaporated, raise the water bath temperature and remove as much acetic acid as possible on the rotary evaporator. Transfer the flask to a high vacuum line for a few hours to remove the residual acetic acid. If the product is still liquid after this process it usually will solidify upon standing.

2. Synthesis of 1,5-hexadiyne. More work needs to be done to develop the procedure for preparation of this compound, but the following procedure is recommended as of this writing.

Quantities of materials should be determined as follows. The amount of 1,2,5,6-tetrabromohexane used should be equimolar with the desired theoretical yield of 1,5-hexadiyne. Since the tetrabromohexane is probably not pure, it might be wise to use an excess over the calculated mass. The mole ratio of tetraoctylammonium bromide to tetrabromohexane should be 1/25, while that for pinacol to tetrabromohexane should be 2/25. In calculating the masses of these materials, remember that the quaternary ammonium salt is 99% and the pinacol is 97%. The amount of potassium hydroxide to be used has been varied as discussed above. We recommend using a high ratio together with a long reaction time to ensure complete reaction. Dehmlo used the equivalent of a 20/1 mole ratio of potassium hydroxide to tetrabromohexane, while Brandsma used 36/1. In calculating the mass, remember that potassium hydroxide is 85%. The amount of solvent should be calculated by adding the masses of all of the above materials and multiplying that sum by 5 mL/g. The reaction works better with a large amount of solvent.

Add freshly machine-powdered potassium hydroxide, tetraoctylammonium bromide, pinacol and most of the solvent to a reaction kettle fitted with an efficient mechanical stirrer and a reflux condenser. *If heat is to be applied, do not add the other two compounds until after the potassium hydroxide has been heated for awhile.* Stir the mixture at about 125 rpm and add the tetrabromohexane slowly. After the addition is completed, add the remaining solvent and increase the stirring rate to 425 rpm. We recommend stirring at room temperature for at least 24 h. The optimum combination of time, temperature, and amount of potassium hydroxide remains to be determined, however. It should be possible to monitor the reaction progress by gas chromatography if quick and easy access to the instrument is available.

When the reaction is complete, filter the reaction mixture, and dry the filtrate over anhydrous magnesium sulfate. After drying, place the liquid on a high vacuum line warm the mixture to 55°C and distill the crude product into a clean liquid-nitrogen-cooled trap. Add ferrous sulfate and *p-tert*-butylcatechol to the crude

product and redistill it.

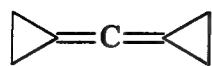
3. Comparison of materials cost for preparations of 1,2,5,6-tetrabromohexane.

<u>Item</u>	<u>Amount per Gram Product</u>	<u>Unit Price</u>	<u>Net Price</u>
sodium bromide (ACS)	1.138 g	\$0.048	\$0.055
sodium perborate	0.851 g	\$0.054	\$0.046
1,5-hexadiene (97%)	0.211 g	\$0.435	\$0.092
<u>glacial acetic acid</u>	<u>8.515 mL</u>	<u>\$0.012</u>	<u>\$0.101</u>
total per gram of product by newer method			\$0.293

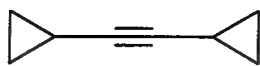
<u>Item</u>	<u>Amount per Gram Product</u>	<u>Unit Price</u>	<u>Net Price</u>
bromine	0.879 g	\$0.105	\$0.093
1,5-hexadiene (97%)	0.211 g	\$0.435	\$0.092
<u>dichloromethane</u>	<u>2.489 mL</u>	<u>\$0.008</u>	<u>\$0.021</u>
total per gram of product by traditional method			\$0.205

In terms of materials costs, a 70% yield by the traditional method is equivalent to a 100 % yield by the newer method. The newer method must be 1.43 times as efficient as the traditional method for materials costs to be equivalent.

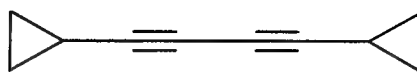
4. Estimated enthalpies of formation for some compounds of interest. Enthalpies of formation were calculated for the compounds whose structures are shown on the next page. Results are shown in the table which follows.



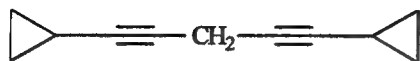
I



II



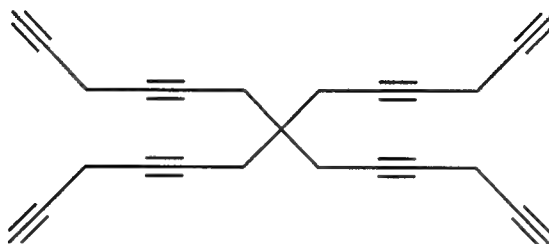
III



IV



V



VI

<u>Compound</u>	ΔH_f	ΔH_f
	(kcal/mol)	(kcal/g)
I	116.6	1.265
II	86.4	0.814
III	137.0	1.050
IV	136.4	0.946
V	199.1	1.183
VI	399.1	1.245

References

1. Shell Development Company, *Rocket Fuels Derivable from Petroleum*, Final Summary Report for Department of the Navy, Bureau of Aeronautics, 1953.
2. Clark, John D., *Ignition! An Informal History of Liquid Rocket Propellants*, Rutgers University Press: New Brunswick, NJ, 1972.
3. Brandsma, L., *Preparative Acetylenic Chemistry*, 2nd ed., Elsevier: Amsterdam, 1988.
4. Peiffer, G., *Bull. Soc. Chim., France*, **1962**, 776.
5. Peiffer, G., *Bull. Soc. Chim., France*, **1962**, 777.
6. Kabalka, G. W., Yang, K., Reddy, N. K. and Narayana, C., *Syn. Comm.*, **1998**, 925.
7. Dehmlow, E. V. and Lissel, M., *Tetrahedron*, **1981**, 37, 1653.
8. Dehmlow, E. V., Thieser, R., Sasson, Y. and Neumann, R., *Tetrahedron*, **1986**, 42, 3569.
9. Benson, S. W. and Buss, J. H., *J. Chem. Phys.*, **1958**, 29, 546.
10. Benson, S. W., *Thermochemical Kinetics*, 2nd ed., Wiley: New York, 1976.
11. Cohen, N., *J. Phys. Chem. Ref. Data*, **1996**, 25, 1411.

**PARTITIONING OF POWER APERATURE PRODUCT
OF
SPACE BASED RADAR**

**Gurnam S. Gill
Department of Electrical and Computer Engineering**

**Naval Postgraduate School
Monterey, CA 93943-5121**

**Final Report for:
Summer Faculty Research Program
Air Force Research Laboratory -Phillips Site**

**Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base,
and
Phillips Site**

September, 1998

**PARTITIONING OF POWER APERATURE PRODUCT
OF
SPACE BASED RADAR**

Gurnam S. Gill

Department of Electrical and Computer Engineering

Naval Postgraduate School

Monterey, CA 93943-5121

Abstract

There may be some operational and cost benefits to be realized by performing space based microwave sensing of surface and air targets with a larger number of smaller satellites than with a fewer number of larger satellites. One way of implementing this concept is by dividing the search space into smaller units and allocating a smaller satellite to each sub space. This is equivalent to partitioning the power aperture product of a large radar into several smaller units. This report analyses power aperture partitioning and quantifies the intuitively expected results. The report concludes that with partitioning there is loss in signal to noise ratio as well as increase in clutter which have to be compensated by increasing the transmitter power or antenna aperture or time on target and additional clutter cancellation to achieve performance equivalent to the unpartitioned case.

PARTITIONING OF POWER APERATURE PRODUCT

OF

SPACE BASED RADAR

Introduction

One of the key points of TechSat 21 program is that it may be cheaper to manufacture, launch and replace smaller satellites than larger satellites. From space based radar (SBR) perspective question arises can performance of one large satellite can be efficiently duplicated by several smaller satellites. In TechSat 21 program the intent is to have cluster of smaller inter linked satellites to form a virtual large satellite. In this approach though the individual satellites may be inexpensive but there in increase in complexity of data exchange between satellites and the signal processing. Furthermore there is fundamental issue whether a radar sensor can be efficiently implemented from a cluster of satellites which jointly form sparse array.

Another approach to replacing large satellite by smaller satellite is to partition search space into smaller pieces and allocating each piece to a smaller satellite. Thus by reducing the search space of each satellite, the size of satellites can be reduced. The transmitter power and antenna aperture primarily determines the size of radar satellite. Questions to be answered are as follows:

- (a) Can power aperture product be partitioned into smaller pieces with out any loss, more specifically, can power P and aperture A of one large satellite be divided into m smaller satellites, each with power P/m and aperture A/m with equivalent search performance?
- (b) What happens to clutter in the partitioning process?

These questions shall be considered for both the air (angular) search and the surface search. Furthermore, more generalized partitioning will be indicated and tradeoffs involved will be pointed.

Air Search

Performance of search radar can be specified as to search solid angle Ω in search time t_s to detect target of radar cross-section σ up to range R_{\max} at given P_d and P_{fa} or given signal to noise ratio (S/N). The equation, which relates these search parameters, is as follows:

$$(1) \quad R_{\max}^4 = \frac{P_{av} A}{F L} \frac{\sigma t_s}{(S/N)_{\min} \Omega} \frac{1}{4\pi k T_o}$$

For constant F , L , $(S/N)_{\min}$, σ above equation can be written as:

$$(2) \quad \Omega = C P_{av} A t_s$$

where $C = \frac{\sigma}{F L (S/N)_{\min} R_{\max}^4 4\pi k T_o}$ is constant.

Equation (2) indicates that search angle Ω is proportional to the product of average transmitter power P_{av} , antenna aperture size A , and search time t_s . This means that these parameters can be traded to cover the same search angle. Trade off between power and aperture is well known. However, this equation also suggests that search time t_s can be traded with power aperture product. By increasing search time t_s , small power aperture product of smaller satellites can be compensated. Dividing both sides of equation (2) by m ,

$$(3) \quad \Omega/m = P_{av}(A/m)t_s$$

Above equation implies that the cover Ω/m aperture size can be reduce by factor of m . To search over angle Ω it would require m smaller apertures of size A/m . However, the power of each smaller aperture is same as that of a single large aperture (search time t_s is assumed constant in this discussion). Above discussion implies that when one large satellite is replaced with m smaller satellites with aperture size A/m , the total power requirement goes up by factor m . This is the inherent disadvantage of partitioning the aperture into smaller apertures. However, if the total power is kept the same in both the cases, the signal to noise ratio will go down by factor of m with partitioning.

The impact of partitioning on clutter can be determined by writing the expression for signal to clutter ratio

$$(4) \quad \frac{s}{c} = \frac{\sigma_t}{\sigma_c} = \frac{\sigma_t}{\sigma^0 (c\tau/2 \sec\Phi) R \theta_A}$$

where σ_t and σ_c are respectively target RCS and clutter radar cross-section, σ^0 is cross section per unit for surface clutter.

Azimuth angle θ_A is given by

$$(5) \quad \theta_A = b \frac{\lambda}{L_A}$$

where L_A is the length of antenna in azimuth dimension. Above two equations indicate that decreasing the aperture increases the clutter. For a square aperture, the clutter will increase by a factor of \sqrt{m} with partitioning. Thus, with partitioning, s/c ratio will decrease by factor of \sqrt{m} which also means for clutter limited situation, the maximum detection range will decrease by a factor of \sqrt{m} .

It should be mentioned that partitioning of aperture could be done in other ways than suggested in the above paragraph by rewriting equation (3) as

$$(6) \quad \Omega/m = c P_{av}(A/m)t_s = c (P_{av}/m)At_s = c P_{av}A (t_s/m)$$

The case suggested by first equality has been discussed earlier. Second equality suggests that it is possible to have m apertures each of aperture size A and power P_{av}/m and the same search time. This is not attractive as it increases the total aperture by factor of m and keeps same total average power and search time because it is costly to launch larger number of large apertures. Similarly third equality suggests that it is possible to reduce search time by increasing total aperture size and average power, which is not an attractive trade off either. The most general partitioning is given by

$$(7) \quad \Omega/m_1 m_2 m_3 = c (P_{av}/m_1) (A/m_2) (t_s/m_3)$$

where m_1 , m_2 , and m_3 can be less than unity, however the product $m_1 m_2 m_3$ should be an integer. This equation suggests $m_1 m_2 m_3$ apertures.

Surface Search Case

Search radar range equation for the surface search need to be derived before making any conclusions. Standard radar range equation for coherent integration of n pulses is given by

$$(8) \quad S/N = \frac{P_t G A \sigma n}{(4\pi)^2 L R^4 k B T_o F}$$

Above equation will be converted into search equation by substituting the expressions for G and n as given below

$$(9) \quad G = \frac{4\pi}{\theta_A \theta_E}$$

where θ_A and θ_E are azimuth and elevation beam widths.

n is obtained as follows

$n = \text{time per beam position} \times \text{PRF}$

$= \frac{\text{Total search time}}{\text{No. of beam positions}} \times \text{PRF}$

$= \frac{\text{Total search time}}{\text{Total area/beam foot print}} \times \text{PRF}$

$= \frac{t_s}{a} \frac{R \theta_A R \theta_E}{\sin \Phi} f_r$

Where " a " is total area to be searched, f_r is pulse repetition frequency and Φ is depression angle of the beam. Substituting expressions for G and n in equation (8) one obtains

$$(10) \quad \frac{S}{N} = \frac{P_{av} A \sigma t_s}{4\pi R^2 L R T_o F a \sin \Phi}$$

which can be written as

$$(11) \quad R_{\max}^2 = \frac{P_{av} A \sigma t_s}{4\pi (S/N)_{\min} L k T_o F a \sin \Phi}$$

Most parameters in the above equation are similar to the air search equation except for the R_{\max} term. The range term in the air search equation has fourth power and in the above equation it is square. This means the radar detection range is more sensitive to the radar parameters for the surface search. Nevertheless, above equation can be written in the same format as equation (2).

$$(12) \quad a = C_1 P_{av} A t_s$$

where a is the area to be searched in time t_s and C_1 is a constant. Thus, all the conclusions drawn for the air search case about the partitioning of power aperture product are also valid for the surface search.

COMPUTATIONAL ASPECTS OF THE SPECTRAL THEORY
OF PHYSICAL AND CHEMICAL BINDING

Robert J. Hinde
Assistant Professor
Department of Chemistry

The University of Tennessee
Knoxville, TN 37996-1600

Final Report for
Summer Faculty Research Program
Air Force Research Laboratory, Propulsion Directorate
Propulsion Sciences and Advanced Concepts Division
Wright Research Site at Edwards AFB

Sponsored by
Air Force Office of Scientific Research
Bolling AFB, Washington DC
and
Wright Research Site

August 1998

COMPUTATIONAL ASPECTS OF THE SPECTRAL THEORY OF PHYSICAL AND CHEMICAL BINDING

Robert J. Hinde
Assistant Professor
Department of Chemistry
The University of Tennessee

Abstract

We develop the computational infrastructure required for spectral theory calculations of the electronic energy of atomic and molecular aggregates in an arbitrary finite spectral product basis set. We test our approach by performing illustrative calculations of the potential surface of Ar_3 clusters in a spectral product basis set composed of four spectral states on each argon atom. We also investigate the completeness of two finite H_2 spectral product basis sets by computing potential energy curves for H_2 using these basis sets. We find that the two basis sets considered here are far from complete, and produce approximate H_2 electronic wavefunctions that exhibit considerable mixed singlet-triplet character.

COMPUTATIONAL ASPECTS OF THE SPECTRAL THEORY OF PHYSICAL AND CHEMICAL BINDING

Robert J. Hinde

1. Introduction

In recent years, considerable effort has been devoted under the auspices of the Air Force Research Laboratory (AFRL) to the development of a "spectral theory" of the physical and chemical binding of molecular aggregates [1-5], in which the molecular electronic Hamiltonian matrix is written *exactly* as a sum of atomic and diatomic "fragment matrices". The great promise of the spectral theory is that these fragment matrices can be computed *once and for all* using conventional *ab initio* electronic structure methods, and then used repeatedly to assemble electronic Hamiltonian matrices for ever-larger polyatomic systems. Such an approach could be used to compute potential energy surfaces for metal atoms embedded in cryogenic matrices, and this goal has provided much of the impetus for in-house AFRL studies of the spectral theory [6-8].

The development of the spectral theory is intimately tied to the notion of an "atomic product" basis set. Such a basis set is generated by associating each electron in a molecule with a specific nucleus, computing the "atomic wavefunctions" for each nucleus and its associated electrons, and then constructing all possible Hartree products of these atomic functions. These Hartree products form a basis set for representing the many-electron wavefunctions of a polyatomic system [3]. Previous applications of the spectral theory [6-8] have been restricted to atomic product basis sets in which only one of the atoms in a system is permitted to contribute excited states to the Hartree product. In this report, we develop the computational infrastructure needed to lift this restriction.

Section 2 begins with a brief review of the spectral theory. We then present three algorithms for assembling the molecular Hamiltonian matrix from atomic and diatomic fragment matrices computed in an *entirely general* (but finite) atomic product basis set. We demonstrate that if a molecule contains two identical nuclei, such a basis set places symmetry-related constraints on the form of one of the fragment matrices used to build the Hamiltonian matrix. Finally, we point out that the Hamiltonian matrix computed using such a basis set is very sparse but also very large, and suggest a truncation of the Hamiltonian matrix related to the truncated configuration interaction matrix familiar from conventional *ab initio* electronic structure calculations.

Section 3 presents some illustrative calculations of the potential surface of Ar_3 using an atomic product basis set in which each argon atom contributes four atomic wavefunctions. These calculations demonstrate the feasibility of implementing the algorithms presented in Sec. 2. Section 4 considers the errors introduced by employing a finite atomic product basis set. The spectral theory is formally exact only in the limit that the atomic product basis set is complete; any finite

atomic product basis set will necessarily be incomplete. We investigate the approach to basis set completeness by performing *ab initio* calculations of H_2 potential curves using two atomic product basis sets of different sizes. We conclude by identifying some areas for further research in Sec. 5.

2. Computation of the spectral theory Hamiltonian matrix

The spectral theory of chemical binding [1-5] is motivated by the observation [9] that the clamped-nucleus Hamiltonian for a molecule can be written exactly as a sum of atomic and diatomic operators. To accomplish this decomposition of the Hamiltonian, we associate each electron in the molecule with a particular nucleus; the collection of (distinguishable) electrons assigned to nucleus α will be denoted $\{\alpha\}$. We then define (using atomic units) the atomic operators

$$\hat{H}^{(\alpha)} = \sum_{i \in \{\alpha\}} \left[-\nabla_i^2 - Z_\alpha / r_{i\alpha} \right] + 1/2 \sum'_{i,j \in \{\alpha\}} 1/r_{ij} \quad (1)$$

and the diatomic operators

$$\hat{V}^{(\alpha,\beta)} = Z_\alpha Z_\beta / R_{\alpha\beta} - \sum_{i \in \{\alpha\}} Z_\beta / r_{i\beta} - \sum_{i \in \{\beta\}} Z_\alpha / r_{i\alpha} + \sum_{\substack{i \in \{\alpha\} \\ j \in \{\beta\}}} 1/r_{ij} \quad (2)$$

where i and j label distinct electrons, α and β label distinct nuclei, and Z_α is the charge of nucleus α . The primed sum in Eq. (1) indicates that terms with $i = j$ are excluded from the sum. The clamped-nucleus Hamiltonian operator for a molecular system with N nuclei is then given by

$$\hat{H} = \sum_{\alpha=1}^N \hat{H}^{(\alpha)} + 1/2 \sum'_{\alpha,\beta=1}^N \hat{V}^{(\alpha,\beta)}. \quad (3)$$

Let

$$\{\phi^{(\alpha)}\} = \{\phi_1^{(\alpha)}(\{\alpha\}), \phi_2^{(\alpha)}(\{\alpha\}), \dots\} \quad (4)$$

be the eigenfunctions (or "spectral states") of the atomic Hamiltonian defined in Eq. (1); the energy eigenvalue associated with $\phi_n^{(\alpha)}$ is $E_n^{(\alpha)}$. The functions $\{\phi^{(\alpha)}\}$ are antisymmetric with respect to the exchange of two electrons in the set $\{\alpha\}$. Now consider the Hartree product of atomic spectral states (or "spectral product")

$$\Phi_{a,b,c,\dots} = \phi_a^{(\alpha)}(\{\alpha\}) \phi_b^{(\beta)}(\{\beta\}) \phi_c^{(\gamma)}(\{\gamma\}) \dots \quad (5)$$

consisting of the product of N functions, one chosen from each set of atomic spectral states [Eq. (4)]. The collection of all possible spectral products forms a complete basis set [3] for representing the many-electron eigenfunctions of the molecular Hamiltonian operator given in Eq. (3). This is true despite the fact that each individual spectral product $\Phi_{a,b,c,\dots}$ is not antisymmetric with

respect to the exchange of two electrons associated with different nuclei, while the many-electron eigenfunctions of \hat{H} are antisymmetric with respect to the exchange of any two electrons in the system [3].

The set of spectral products $\{\Phi_{a,b,c,\dots}\}$ is infinite, because each atomic Hamiltonian $\hat{H}^{(\alpha)}$ has an infinite number of eigenfunctions. However, if we truncate the atomic spectral state sets [Eq. (4)], the set of possible spectral products becomes finite, and we can use this set of functions as a computational basis for approximating the molecular electronic wavefunctions. Then we are faced with the task of evaluating Hamiltonian matrix elements such as $\langle \Phi_{a,b,c,\dots} | \hat{H} | \Phi_{f,g,h,\dots} \rangle$. (We note in passing that truncation of the spectral product set renders it incomplete, and convergence tests must be performed to insure that the truncated basis set is sufficiently large to represent the eigenfunctions of \hat{H} reasonably well. We will return to this point below.)

A. Computation of Hamiltonian matrix elements

To construct the Hamiltonian matrix \mathbf{H} in this truncated atomic product basis, we must first develop a scheme for ordering the basis functions $\{\Phi_{a,b,c,\dots}\}$. We have adopted the following ordering scheme. Choose some convention for ordering the atomic eigenfunctions (or atomic spectral states) associated with each nucleus. Then order the spectral products so that atomic state indices for atoms farther to the right in Eq. (5) run to completion before those indices for atoms farther to the left in Eq. (5). The compound subscript a, b, c, \dots which appears in Eq. (5) then cycles through all possible values in a fashion reminiscent of an automobile odometer.

To evaluate Hamiltonian matrix elements in the spectral product basis, it is useful to have an algorithm for converting the compound subscript a, b, c, \dots which identifies a basis function into the sequence number of this basis function in the ordered list of spectral products. Here we present such an algorithm. Define $n(\alpha)$ to be the number of atomic eigenfunctions in the (truncated) set of spectral states associated with nucleus α . Suppose we are given the compound subscript a, b, c, \dots associated with a particular spectral product. Let $s(\alpha)$ be the spectral state index for nucleus α in this compound subscript. From Eq. (5) we see that $s(\alpha) = a$, $s(\beta) = b$, and $s(\gamma) = c$ in this example. To compute the sequence number z of this spectral product in the ordered set of spectral products, we execute the following steps.

Algorithm I:

```

 $z := 0;$ 
for  $1 \leq p \leq N$ 
     $z := z \times n(p) + s(p) - 1;$ 
 $z := z + 1;$ 

```

The following algorithm reverses this process; given the sequence number z of an spectral product function, it extracts the list of atomic spectral state indices from right to left.

Algorithm II:

```

 $k := z;$ 
for  $N \geq p \geq 1$ 
{
 $s(p) := 1 + [(k - 1) \bmod n(p)];$ 
 $k := 1 + (k - 1)/n(p);$ 
}

```

Algorithm II employs FORTRAN integer arithmetic, in which $5/3 = 1$ and $5 \bmod 3 = 2$. Using these two algorithms, we can identify matrix elements of the form $\langle \Phi_{a,b,c,\dots} | \hat{H} | \Phi_{f,g,h,\dots} \rangle$ with specific row-column entries of the Hamiltonian matrix \mathbf{H} .

Now we address the problem of evaluating a specific element of \mathbf{H} . Suppose that the spectral products $\Phi_{a,b,c,\dots}$ and $\Phi_{f,g,h,\dots}$ have respective sequence numbers y and z . We seek to compute the matrix element $H_{yz} = \langle \Phi_{a,b,c,\dots} | \hat{H} | \Phi_{f,g,h,\dots} \rangle$. From Eq. (3), we see that H_{yz} is a sum of matrix elements of the atomic and diatomic operators $\hat{H}^{(\alpha)}$ and $\hat{V}^{(\alpha,\beta)}$; we now turn our attention to these matrix elements.

The matrix element $H_{yz}^{(\alpha)}$ is given by

$$H_{yz}^{(\alpha)} = \Delta_{yz}(\alpha) \Delta_{yz}(\beta) \Delta_{yz}(\gamma) \cdots \Delta_{yz}(N) E_{s_z(\alpha)}^{(\alpha)} \quad (6)$$

where $s_z(\alpha)$ denotes the spectral state index for nucleus α in the spectral product with sequence number z , and $\Delta_{yz}(\alpha)$ is a Kronecker delta function evaluated over atomic spectral state indices: $\Delta_{yz}(\alpha) = \delta[s_y(\alpha), s_z(\alpha)]$. The Kronecker delta function appears in Eq. (6) as a consequence of the orthonormality of the atomic eigenfunctions listed in Eq. (4). The matrix element $V_{yz}^{(\alpha,\beta)}$ is given by

$$V_{yz}^{(\alpha,\beta)} = \Delta_{yz}(\gamma) \cdots \Delta_{yz}(N) \langle \phi_{s_y(\alpha)}^{(\alpha)} \phi_{s_y(\beta)}^{(\beta)} | \hat{V}^{(\alpha,\beta)} | \phi_{s_z(\alpha)}^{(\alpha)} \phi_{s_z(\beta)}^{(\beta)} \rangle. \quad (7)$$

For notational convenience, we will define $v_{yz}^{(\alpha,\beta)} = \langle \phi_{s_y(\alpha)}^{(\alpha)} \phi_{s_y(\beta)}^{(\beta)} | \hat{V}^{(\alpha,\beta)} | \phi_{s_z(\alpha)}^{(\alpha)} \phi_{s_z(\beta)}^{(\beta)} \rangle$.

Next, consider the operator

$$\hat{W}^{(\alpha,\beta)} = \hat{H}^{(\alpha)} + \hat{H}^{(\beta)} + \hat{V}^{(\alpha,\beta)}. \quad (9)$$

This operator is the clamped-nucleus Hamiltonian for the diatomic molecule composed of nuclei α and β and their associated electrons $\{\alpha\}$ and $\{\beta\}$. The eigenfunctions of $\hat{W}^{(\alpha,\beta)}$ are simply the electronic wavefunctions of this diatomic molecule, which satisfy the diatomic Schrödinger equation

$$\hat{W}^{(\alpha,\beta)} \Psi_w^{(\alpha,\beta)}(R_{\alpha\beta}) = E_w^{(\alpha,\beta)}(R_{\alpha\beta}) \Psi_w^{(\alpha,\beta)}(R_{\alpha\beta}). \quad (10)$$

The index w distinguishes different electronic states of the diatomic molecule. These wavefunctions can be approximated using conventional *ab initio* electronic structure techniques. In order to exploit the cylindrical symmetry of the diatomic molecule, such a calculation will typically employ a molecule-fixed coordinate system in which the z axis is aligned with the internuclear vector $\mathbf{R}_{\alpha\beta}$.

The matrix element $V_{yz}^{(\alpha,\beta)}$ can be related to the set of diatomic Born–Oppenheimer potential energy curves $\{E_w^{(\alpha,\beta)}(R_{\alpha\beta})\}$ as follows. From Eq. (9), we have

$$\hat{V}^{(\alpha,\beta)} = \hat{W}^{(\alpha,\beta)} - \hat{H}^{(\alpha)} - \hat{H}^{(\beta)}, \quad (11)$$

so that

$$v_{yz}^{(\alpha,\beta)} = \langle \phi_{s_y(\alpha)}^{(\alpha)} \phi_{s_y(\beta)}^{(\beta)} | \hat{W}^{(\alpha,\beta)} | \phi_{s_z(\alpha)}^{(\alpha)} \phi_{s_z(\beta)}^{(\beta)} \rangle - \Delta_{yz}(\alpha) \Delta_{yz}(\beta) (E_{s_z(\alpha)}^{(\alpha)} + E_{s_z(\beta)}^{(\beta)}) \quad (12)$$

and the problem of computing $V_{yz}^{(\alpha,\beta)}$ reduces to that of computing matrix elements of the operator $\hat{W}^{(\alpha,\beta)}$ in the basis set of atomic spectral state products centered on nuclei α and β .

The matrix elements of $\hat{W}^{(\alpha,\beta)}$ in the basis set of diatomic wavefunctions $\{\Psi_w^{(\alpha,\beta)}(R_{\alpha\beta})\}$ are known; these are simply the Born–Oppenheimer potential curves $\{E_w^{(\alpha,\beta)}(R_{\alpha\beta})\}$. Hence the desired matrix elements of $\hat{W}^{(\alpha,\beta)}$ in the atomic spectral state product basis set can be approximated in terms of a transformation between this basis set and the diatomic wavefunction basis set:

$$\begin{aligned} \langle \phi_{s_y(\alpha)}^{(\alpha)} \phi_{s_y(\beta)}^{(\beta)} | \hat{W}^{(\alpha,\beta)} | \phi_{s_z(\alpha)}^{(\alpha)} \phi_{s_z(\beta)}^{(\beta)} \rangle \approx \\ \sum_w \langle \phi_{s_y(\alpha)}^{(\alpha)} \phi_{s_y(\beta)}^{(\beta)} | \Psi_w^{(\alpha,\beta)}(R_{\alpha\beta}) \rangle E_w^{(\alpha,\beta)}(R_{\alpha\beta}) \langle \Psi_w^{(\alpha,\beta)}(R_{\alpha\beta}) | \phi_{s_z(\alpha)}^{(\alpha)} \phi_{s_z(\beta)}^{(\beta)} \rangle. \end{aligned} \quad (13)$$

This is an approximation because (1) both the spectral product and diatomic wavefunction basis sets are typically finite (and hence incomplete), and (2) the diatomic wavefunctions and potential curves used in Eq. (13) are typically only approximate solutions to Eq. (10). Note that we have used the fact that the diatomic wavefunctions are orthonormal in obtaining the right hand side of Eq. (13).

Finally, we point out that Eq. (13) implies that the atomic spectral states of Eq. (4) and the diatomic wavefunctions of Eq. (10) are defined in the same coordinate system. In general this will not be the case. As we noted above, the diatomic wavefunctions of Eq. (10) will typically be computed in a molecule-fixed coordinate system tied to the internuclear vector $\mathbf{R}_{\alpha\beta}$. The atomic spectral states of Eq. (4), however, will usually be defined in a laboratory-fixed coordinate system. Hence we must transform the laboratory-fixed atomic spectral states of Eq. (4) into the molecule-fixed coordinate system used to solve Eq. (10). This can be accomplished using a direct product of rotation matrices as described in Ref. [3]; we will denote the appropriate rotation matrix product by $\mathbf{D}^{(\alpha,\beta)}(\hat{\mathbf{R}}_{\alpha\beta})$.

Suppose we define the transformation matrix $U^{(\alpha,\beta)}(R_{\alpha\beta})$ which has as elements the “overlap integrals”

$$U_{wz}^{(\alpha,\beta)} = \langle \Psi_w^{(\alpha,\beta)}(R_{\alpha\beta}) | \phi_{s_z(\alpha)}^{(\alpha)} \phi_{s_z(\beta)}^{(\beta)} \rangle, \quad (14)$$

where both the atomic spectral states and the diatomic wavefunction are defined in a coordinate system in which the z axis is aligned with $\mathbf{R}_{\alpha\beta}$. Then Eq. (13) becomes

$$\langle \phi_{s_y(\alpha)}^{(\alpha)} \phi_{s_y(\beta)}^{(\beta)} | \hat{W}^{(\alpha,\beta)} | \phi_{s_z(\alpha)}^{(\alpha)} \phi_{s_z(\beta)}^{(\beta)} \rangle \approx \left[\mathbf{D}^{(\alpha,\beta)}(\hat{\mathbf{R}}_{\alpha\beta})^\dagger \times \mathbf{U}^{(\alpha,\beta)}(R_{\alpha\beta})^\dagger \times \mathbf{E}^{(\alpha,\beta)}(R_{\alpha\beta}) \times \mathbf{U}^{(\alpha,\beta)}(R_{\alpha\beta}) \times \mathbf{D}^{(\alpha,\beta)}(\hat{\mathbf{R}}_{\alpha\beta}) \right]_{yz} \quad (15)$$

where $\mathbf{E}^{(\alpha,\beta)}(R_{\alpha\beta})$ is the diagonal matrix of Born–Oppenheimer potential energy curves obtained from Eq. (10).

We now pause for a moment to consider the sizes of the various matrices we have introduced. The number of distinct atomic products, and the size of the matrices \mathbf{H} and \mathbf{V} defined in Eqs. (6) and (7), is

$$P = \prod_{\alpha=1}^N n(\alpha). \quad (16)$$

The matrices \mathbf{v} and \mathbf{W} defined in Eqs. (12) and (13) could also be thought of as $P \times P$ matrices; however, in such an approach many of the elements of these matrices are redundant. To illustrate this, we consider a triatomic system in which each atom has two spectral states. Then

$$v_{3,5}^{(1,2)} = \langle \phi_1^{(1)} \phi_2^{(2)} \phi_1^{(3)} | \hat{V}^{(1,2)} | \phi_2^{(1)} \phi_1^{(2)} \phi_1^{(3)} \rangle = \langle \phi_1^{(1)} \phi_2^{(2)} \phi_2^{(3)} | \hat{V}^{(1,2)} | \phi_2^{(1)} \phi_1^{(2)} \phi_2^{(3)} \rangle = v_{4,6}^{(1,2)}. \quad (17)$$

Note that we have used the “odometer” ordering scheme outlined above to order the spectral product basis functions.

A more efficient approach is to consider the matrices $\mathbf{v}^{(\alpha,\beta)}$ and $\mathbf{W}^{(\alpha,\beta)}$ to be $Q \times Q$ matrices, where $Q = n(\alpha)n(\beta)$. These smaller matrices can be thought of as matrix representations of the corresponding operators $\hat{V}^{(\alpha,\beta)}$ and $\hat{W}^{(\alpha,\beta)}$ in a “reduced” atomic spectral state product basis set, consisting of products of one spectral state associated with nucleus α and one spectral state associated with nucleus β . The price we must pay for this simplification is that we must develop a *second* ordering scheme for this reduced basis set, along with the bookkeeping needed to associate matrix elements in the reduced basis set with those in the original basis set. In addition, the size of $\mathbf{v}^{(\alpha,\beta)}$ and $\mathbf{W}^{(\alpha,\beta)}$ now depends on the identities of the nuclei α and β .

For consistency, we order the reduced basis set in exactly the same manner as the larger spectral product basis set, using the “odometer” sequence developed earlier. In our triatomic model system, then, the diatomic spectral products $\phi_1^{(1)} \phi_2^{(2)}$ and $\phi_2^{(1)} \phi_1^{(2)}$ have respective sequence numbers 2 and 3 in the reduced spectral product basis set associated with nuclei 1 and 2.

We can now outline an algorithm for computing the Hamiltonian matrix \mathbf{H} in the atomic spectral product basis. The algorithm consists of two phases. In the first phase, the diatomic $\mathbf{v}^{(\alpha,\beta)}$ matrices are computed and stored. The Hamiltonian matrix is then assembled from these small matrices in the second phase.

Algorithm III:

```

for  $1 \leq \alpha \leq N - 1$ 
  for  $\alpha + 1 \leq \beta \leq N$ 
    for  $1 \leq i \leq n(\alpha) \times n(\beta)$ 
      for  $1 \leq j \leq n(\alpha) \times n(\beta)$ 
        {
          compute the atomic spectral state indices  $\{s_i(\alpha), s_i(\beta)\}$  and  $\{s_j(\alpha), s_j(\beta)\}$ 
            from  $i$  and  $j$  using Algorithm II;
          compute  $[\mathbf{v}^{(\alpha,\beta)}]_{ij}$  using Eqs. (12) and (15);
        }
  for  $1 \leq y \leq P$ 
    for  $y \leq z \leq P$ 
      {
         $H_{yz} := 0$ ;
        compute the atomic spectral state indices  $\{s_y\}$  and  $\{s_z\}$  using Algorithm II;
        if  $y = z$  then
          for  $1 \leq \alpha \leq N$ 
             $H_{yz} := H_{yz} + E_{s_z(\alpha)}^{(\alpha)}$ ;
          for  $1 \leq \alpha \leq N - 1$ 
            for  $\alpha + 1 \leq \beta \leq N$ 
              if  $\Delta_{yz}(\gamma) = 1$  for all  $\gamma \neq \alpha$  and  $\gamma \neq \beta$  then
                {
                  use Algorithm I to convert  $\{s_y(\alpha), s_y(\beta)\}$  and  $\{s_z(\alpha), s_z(\beta)\}$  into "reduced"
                    spectral state sequence numbers  $i$  and  $j$ ;
                   $H_{yz} := H_{yz} + [\mathbf{v}^{(\alpha,\beta)}]_{ij}$ ;
                }
              if  $y \neq z$  then  $H_{zy} := H_{yz}$ ;
            }
      }

```

A FORTRAN program which implements this algorithm is available from the author upon request.

B. Constraints on the \mathbf{U} matrix for a homonuclear diatomic molecule

Suppose that two identical nuclei α and β are located on the z axis such that the rotation matrix product $\mathbf{D}^{(\alpha,\beta)}(\hat{\mathbf{R}}_{\alpha\beta}) = \mathbf{I}$. Then the matrix which appears on the right hand side of Eq. (15) becomes

$$\mathbf{U}^{(\alpha,\beta)}(R_{\alpha\beta})^\dagger \times \mathbf{E}^{(\alpha,\beta)}(R_{\alpha\beta}) \times \mathbf{U}^{(\alpha,\beta)}(R_{\alpha\beta}). \quad (18)$$

Now consider the matrix elements (evaluated in the atomic spectral product basis)

$$\langle ab | \hat{W}^{(\alpha,\beta)} | cd \rangle = \langle \phi_a^{(\alpha)} \phi_b^{(\beta)} | \hat{W}^{(\alpha,\beta)} | \phi_c^{(\alpha)} \phi_d^{(\beta)} \rangle \quad (19)$$

and

$$\langle ba | \hat{W}^{(\alpha,\beta)} | dc \rangle = \langle \phi_b^{(\alpha)} \phi_a^{(\beta)} | \hat{W}^{(\alpha,\beta)} | \phi_d^{(\alpha)} \phi_c^{(\beta)} \rangle. \quad (20)$$

Because α and β are identical nuclei, these two matrix elements must be invariant to within a sign change. Specifically,

$$\langle ab | \hat{W}^{(\alpha,\beta)} | cd \rangle = (-1)^p \langle ba | \hat{W}^{(\alpha,\beta)} | dc \rangle \quad (21)$$

where

$$p = \sigma(\phi_a^{(\alpha)}) + \sigma(\phi_b^{(\beta)}) + \sigma(\phi_c^{(\alpha)}) + \sigma(\phi_d^{(\beta)}) \quad (22)$$

and $\sigma(\phi)$ indicates the symmetry properties of the atomic eigenfunction ϕ under the transformation $(x, y, z) \rightarrow (x, y, -z)$. If ϕ is invariant under this transformation, then $\sigma(\phi) = 0$; if ϕ changes sign under this transformation, then $\sigma(\phi) = 1$. Eq. (21) places constraints on the \mathbf{U} matrices of homonuclear diatomic molecules.

Let i and j be the respective sequence numbers of the functions $\phi_a^{(\alpha)} \phi_b^{(\beta)}$ and $\phi_c^{(\alpha)} \phi_d^{(\beta)}$ in the reduced spectral product basis set associated with nuclei α and β . Similarly, let i' and j' be the respective sequence numbers of the functions $\phi_b^{(\alpha)} \phi_a^{(\beta)}$ and $\phi_d^{(\alpha)} \phi_c^{(\beta)}$ in this basis set. Then Eq. (21) becomes

$$\left[\mathbf{U}^{(\alpha,\beta)}(R_{\alpha\beta})^\dagger \times \mathbf{E}^{(\alpha,\beta)}(R_{\alpha\beta}) \times \mathbf{U}^{(\alpha,\beta)}(R_{\alpha\beta}) \right]_{ij} = (-1)^p \left[\mathbf{U}^{(\alpha,\beta)}(R_{\alpha\beta})^\dagger \times \mathbf{E}^{(\alpha,\beta)}(R_{\alpha\beta}) \times \mathbf{U}^{(\alpha,\beta)}(R_{\alpha\beta}) \right]_{i'j'}, \quad (23)$$

or, recalling that $\mathbf{E}^{(\alpha,\beta)}(R_{\alpha\beta})$ is diagonal (and omitting the $R_{\alpha\beta}$ dependence of the matrix elements for brevity),

$$\sum_w U_{wi}^* E_w U_{wj} = (-1)^p \sum_w U_{wi'}^* E_w U_{wj'}. \quad (24)$$

Rewriting Eq. (24) as

$$\sum_w E_w \left[U_{wi}^* U_{wj} - (-1)^p U_{wi'}^* U_{wj'} \right] = 0, \quad (25)$$

we find that if

$$U_{wi}^* U_{wj} = (-1)^P U_{wi'}^* U_{wj'} \quad (26)$$

for each diatomic molecular eigenstate w , then Eq. (21) is satisfied.

C. \mathbf{H} is a sparse matrix

The size of the Hamiltonian matrix, P , is given by Eq. (16). Whenever each atom has two or more spectral states, P grows exponentially with N ; for the special case of N identical atoms with w spectral states each, $P = w^N$. Hence as the number of atoms in a system increases, \mathbf{H} rapidly becomes a very large matrix. However, \mathbf{H} is also very sparse.

Consider row y of \mathbf{H} . Eqs. (6) and (7) show that the only nonzero elements H_{yz} in this row are those for which the corresponding spectral products $\Phi_{a,b,c,\dots}$ and $\Phi_{f,g,h,\dots}$ differ in zero, one, or two positions. The number of spectral products which differ from $\Phi_{a,b,c,\dots}$ in exactly one position is $\sum_{\alpha=1}^N [s(\alpha) - 1]$. The number of spectral products which differ from $\Phi_{a,b,c,\dots}$ in exactly two positions is

$$1/2 \sum_{\alpha, \beta=1}^N [s(\alpha) - 1][s(\beta) - 1] = 1/2 \left[\left(\sum_{\alpha=1}^N [s(\alpha) - 1] \right)^2 - \sum_{\alpha=1}^N [s(\alpha) - 1]^2 \right]. \quad (27)$$

The total number of nonzero elements in row y of \mathbf{H} is therefore

$$P^* = 1 + \sum_{\alpha=1}^N [s(\alpha) - 1] + 1/2 \left[\left(\sum_{\alpha=1}^N [s(\alpha) - 1] \right)^2 - \sum_{\alpha=1}^N [s(\alpha) - 1]^2 \right], \quad (28)$$

which is $O(N^2)$. Although the total number of elements in row y of \mathbf{H} grows exponentially with N , the number of nonzero elements in this row grows only polynomially with N . Therefore \mathbf{H} is sparse, and grows sparser as N increases.

Nevertheless, the exponential growth in the size of \mathbf{H} with N is a matter of some concern. Suppose we decide to compute the lowest eigenvalue E_0 of \mathbf{H} using a diagonalization scheme for large sparse matrices. We can avoid storing the entire $P \times P$ Hamiltonian matrix in memory by using Algorithm III presented above to compute the (nonzero) elements of \mathbf{H} one at a time as they are needed in the diagonalization process. However, we must have at least enough memory (in the form of either RAM or disk) to store a single vector of dimension P , which represents the approximate eigenvector of \mathbf{H} corresponding to the eigenvalue E_0 . For a system of 40 atoms with only two spectral states each, we must store $2^{40} \approx 1.1 \times 10^{12}$ floating point numbers; if double-precision arithmetic is used, this will require 8 TB of storage! Hence although \mathbf{H} is in fact sparse, computation of the lowest eigenvalue of \mathbf{H} is unfeasible for systems of even moderate size.

A similar problem is encountered in *ab initio* configuration interaction calculations. The full CI matrix for a molecular system grows unmanageably large as the number of electrons in the system increases; consequently, only truncated CI calculations are feasible for large molecules, advances in direct CI techniques [10] notwithstanding. Truncation of the spectral theory Hamiltonian matrix may prove useful for practical (but approximate) calculations of the lowest eigenvalue(s) of \mathbf{H} .

For instance, we might consider a "singles and doubles" truncation of \mathbf{H} defined as follows. Let Φ_0 be the spectral product formed by choosing the lowest energy atomic eigenfunction from each atomic spectral set defined by Eq. (4). (We tacitly assume that the ground state of each atomic Hamiltonian operator is nondegenerate.) Select those functions from the spectral product basis $\{\Phi_{a,b,c,\dots}\}$ which differ from Φ_0 by at most two atomic functions. This can be done easily using Algorithm III. Finally, construct the truncated Hamiltonian matrix \mathbf{H}_{sd} (which is a dense $P^* \times P^*$ matrix) using this reduced basis set and compute the lowest eigenvalue of \mathbf{H}_{sd} .

A similar approach could be used to include the effects of "triple" and higher excitations in the spectral product basis set by applying Algorithm III recursively. We might anticipate that a good approximation to the lowest eigenvalue E_0 of the full Hamiltonian matrix can be obtained by truncating the matrix at a relatively low level of excitation, particularly if the excited states of the atomic Hamiltonians [Eq. (3)] are substantially higher in energy than the ground states. However, computational tests of this hypothesis are required.

3. Illustrative calculations for Ar_3

Here we present some preliminary studies of the potential surfaces of Ar_3 within the spectral theory framework. We employ a model in which each argon atom has four atomic spectral states [Eq. (4)]: the 1S_0 ground state and the triply degenerate 1P_1 excited state. We will represent the three degenerate excited states by the term symbols $^1P_{-1}$, 1P_0 , and $^1P_{+1}$; the subscript in these symbols can be thought of as the M_J quantum number. (Note that 1P_1 and $^1P_{+1}$ represent different entities.) The energy of the 1S_0 ground state is taken to be zero; the energy of the 1P_1 state is then $E^* = 95399 \text{ cm}^{-1}$ [11]. The ordering of the atomic spectral states is: $\{^1S_0, ^1P_{-1}, ^1P_0, ^1P_{+1}\}$.

In this model, the Hamiltonian matrix \mathbf{H} for Ar_3 is a 64×64 matrix, while the diatomic matrices $\mathbf{E}(R)$, $\mathbf{U}(R)$, and $\mathbf{v}(R)$ are 16×16 matrices. The Born-Oppenheimer Ar_2 potential curves which are the diagonal elements of \mathbf{E} are defined as follows. The potential curve which correlates with the $^1S_0 + ^1S_0$ limit as $R \rightarrow \infty$ is taken to be the empirical $\text{Ar}_2 X^1\Sigma_g^+$ potential of Aziz and Slaman [12]; this is $E_1(R)$ in Eq. (13). The six potential curves which correlate with the $^1S_0 + ^1P_1$ asymptotic limit (E_2, E_3, E_4, E_5, E_9 , and E_{13}) are assumed to be independent of R , and are set to $V(R) = E^*$. (Our model could be refined by using *ab initio* curves for these $^1\Sigma$ and $^1\Pi$ states of Ar_2 [13].) The nine remaining diatomic potential curves, which correlate with the $^1P_1 + ^1P_1$ asymptotic limit, are also assumed to be independent of R and are set to $V(R) = 2E^*$.

We employ an empirical \mathbf{U} matrix. The following elements of \mathbf{U} depend on a parameter $\gamma(R)$ which controls the strength of the non-pairwise-additive contribution to the Ar_n potential energy:

$$\begin{aligned} U_{1,1} &= \cos \gamma(R) & U_{1,3} &= -U_{1,9} = (1/\sqrt{2}) \sin \gamma(R) \\ U_{3,1} &= -\sin \gamma(R) & U_{3,3} &= -U_{3,9} = (1/\sqrt{2}) \cos \gamma(R) \\ & & U_{9,3} &= U_{9,9} = 1/\sqrt{2}. \end{aligned} \quad (29)$$

Rows 1, 3, and 9 of \mathbf{U} represent, respectively, the $X^1\Sigma_g^+$ curve, the first excited $1\Sigma_g^+$ curve, and the lowest $1\Sigma_u$ curve of Ar_2 . The elements of \mathbf{U} which are not listed in Eq. (29) are set to $U_{ij} = \delta(i, j)$. This \mathbf{U} matrix is Hermitian and satisfies Eq. (26). We set $\gamma(R) = \gamma_0 \exp(-R/R_{\text{eq}})$ where γ_0 is a dimensionless free parameter which controls the strength of the non-pairwise forces, and $R_{\text{eq}} = 3.761 \text{ \AA}$ is the equilibrium bond length of the Aziz-Slaman Ar_2 potential curve.

We now present calculations of the ground state potential surface for Ar_3 in C_{2v} and C_s geometries. First we consider a C_{2v} Ar_3 cluster, with fixed $R_{12} = R_{23} = R_{\text{eq}}$ and variable θ_{123} . Figure 1 shows how the potential energy of this Ar_3 cluster depends on the angle θ_{123} and on the parameter γ_0 . (For $\gamma_0 = 0$, the potential computed using the spectral theory reduces to the pairwise-additive limit.) The non-pairwise forces which emerge from our spectral theory model stabilize near-equilateral Ar_3 configurations and destabilize near-collinear configurations.

Next we consider the fragmentation of Ar_3 to give $\text{Ar}_2 + \text{Ar}$. We fix $R_{12} = R_{\text{eq}}$ and $\theta_{123} = 90^\circ$, and increase R_{23} from 3.4 \AA to 16 \AA . Figure 2 shows how the potential energy along this pathway depends on the strength parameter γ_0 . Note that as $R_{23} \rightarrow \infty$ all potential surfaces approach the pairwise additive limit, as they must. Evidently the non-pairwise forces in our spectral theory model are fairly long-range, at least for the two γ_0 values we have considered.

In view of the approximations made in these calculations—namely, the use of an empirical \mathbf{U} matrix and the *ad hoc* nature of the excited state Ar_2 potential curves—the results depicted in Figs. 1 and 2 are largely illustrative. Nevertheless, the present calculations demonstrate that spectral theory calculations of the potential surfaces of small molecules are feasible even in an spectral product basis set which includes several spectral states for all atoms. The results presented in Figs. 1 and 2 are some of the first calculations using such a basis set.

4. Completeness of the truncated spectral product basis: *ab initio* tests for H_2

In Sec. 2, we noted that truncation of the atomic eigenfunction sets renders the spectral product basis set incomplete. Here we investigate the degree of completeness of a finite spectral product basis set for the H_2 system by computing Born–Oppenheimer potential curves for H_2 in the truncated spectral product basis. Our approach is quite straightforward: we simply evaluate the elements of the matrix $\mathbf{W}(R)$, which represents the diatomic Hamiltonian operator $\hat{W}^{(\alpha, \beta)}$ [Eq. (9)] in the spectral product basis set. The eigenvalues of $\mathbf{W}(R)$ are the potential curves we seek. As

the spectral product basis set approaches completeness, these curves should approach the exact Born–Oppenheimer potential curves for H_2 .

The atomic spectral states for hydrogen are one-particle atomic orbitals. If we choose to represent these orbitals as linear combinations of Gaussian-type orbitals (GTOs), we can use conventional quantum chemistry packages to calculate the one- and two-electron integrals required to assemble \mathbf{W} ; this substantially reduces the computational effort needed to obtain the desired potential curves. We used Gaussian 94 [14] to calculate these one- and two-electron integrals. Unfortunately, this approach is not readily extended to other diatomic molecules. However, studies of H_2 are important not only for the fundamental insight into the spectral theory which they provide, but also because H_2 is a major component of several hypothetical high energy density fuels.

We computed \mathbf{W} using two different spectral product basis sets. In the first basis set, each hydrogen atom has five spectral states, which are contracted GTO representations of the 1s, 2s, and triply degenerate 2p orbitals. The exponents and contraction coefficients for these GTO representations are listed in Table I, along with the corresponding atomic energies. The five-GTO representations of the 1s and 2p orbitals are taken from Ref. [15]; the GTO expansion of the 2s orbital was computed by uncontracting the GTO expansions for the 1s and 2p orbitals and performing a full CI calculation of the hydrogen atom in this ten-GTO basis set. We will refer to the spectral product basis set constructed from these 1s, 2s, and 2p orbitals as the “small” basis set; it has $5 \times 5 = 25$ spectral product functions.

In the “large” spectral product basis set, each hydrogen atom has 14 spectral states. These are the five spectral states listed above (the 1s, 2s, and 2p orbitals) plus a set of diffuse s, p, and d orbitals. The diffuse orbitals used to augment the five original spectral states were determined by performing a second full CI calculation of the hydrogen atom in a basis set consisting of the contracted GTO expansions in Table I and the diffuse GTO expansions in Table II. Because the expansions listed in Table I are fairly good representations of the hydrogen atom’s 1s, 2s, and 2p orbitals, they emerge virtually unchanged from the CI calculation as natural orbitals of the hydrogen atom. The natural orbitals corresponding to the third lowest s state, the second lowest p state, and the lowest d state of the CI calculation are the diffuse orbitals used to augment the original 1s, 2s, and 2p orbitals; they are listed in Table III. Our construction of these diffuse orbitals insures that the collection of 14 atomic spectral states forms an orthonormal set. The large spectral product basis set includes $14 \times 14 = 196$ spectral product functions.

In Fig. 3 we show the lowest energy potential curves obtained by diagonalizing the \mathbf{W} matrix in the small and large spectral product basis sets. We also show in this figure the exact Born–Oppenheimer $X^1\Sigma_g^+$ and $b^3\Sigma_u^+$ curves [16] for H_2 . The curves computed using the spectral product basis set are qualitatively correct—they predict that H_2 is a stable molecule—but they are poor quantitative predictions of the H_2 potential. The predicted H_2 binding energy is much too low, and

the equilibrium H_2 bond length is too large. These results indicate that neither spectral product basis set is near completion. The curve computed using the large spectral product basis set is marginally better than that computed with the small basis set.

Insight into the deficiencies of the spectral product basis sets is provided by Fig. 4. Here, we plot the expectation value of the electron exchange operator \hat{P}_{12} evaluated over the spatial part of the two approximate $X^1\Sigma_g^+$ electronic wavefunctions. For a pure singlet state, $\langle \hat{P}_{12} \rangle = 1$; for a pure triplet state, $\langle \hat{P}_{12} \rangle = -1$. The approximate electronic wavefunctions computed using the spectral product basis sets are neither singlets nor triplets; rather, they have considerable "singlet-triplet average" character. Near their respective potential minima, both wavefunctions are apparently a roughly 3:1 mix of singlet and triplet character, giving $\langle \hat{P}_{12} \rangle \approx 0.5$. The wavefunction computed using the large basis set is slightly closer to a pure singlet, in accord with the results presented in Fig. 3.

5. Summary and future studies

In this report, we have addressed several of the computational issues which arise in the spectral theory of physical and chemical binding when a general finite set of atomic spectral states is used to construct the spectral product basis set. We presented three algorithms which can be used to construct the Hamiltonian matrix in such a basis set, and demonstrated their feasibility by performing illustrative spectral theory calculations of the Ar_3 potential surface.

Two important points must be addressed in future studies of the spectral theory. First, computations must be performed which test the hypothesis advanced in Sec. 2 that accurate electronic energies can be obtained using spectral theory Hamiltonian matrices truncated at low levels of excitation. If this hypothesis is correct, then energy calculations for large molecules will be possible using the spectral theory. Second, the preliminary studies of basis set completeness presented in Sec. 4 should be extended to substantially larger basis sets in order to confirm the theorem [3] that the spectral product basis set can represent many-electron wavefunctions, which are fully antisymmetric with respect to the exchange of any two electrons, as a sum of functions which do not possess such symmetry properties. Work along these lines is in progress.

Acknowledgments

This work was supported by the Air Force Office of Scientific Research through the award of a Summer Faculty Research Grant administered by Research and Development Laboratories. I would like to thank the staff of AFRL/PRSP, Edwards AFB for their gracious hospitality during my stay at AFRL and for many helpful conversations and suggestions.

References

- [1] P. W. Langhoff, J. Boatz, and M. E. Fajardo, in *Proceedings of the High Energy Density Matter (HEDM) Contractors' Conference*, T. L. Thompson and S. L. Rodgers, eds., PL-TR-94-3036 (1994).
- [2] P. W. Langhoff, in *Proceedings of the High Energy Density Matter (HEDM) Contractors' Conference*, P. G. Carrick and S. Tam, eds., PL-TR-95-3039 (1996).
- [3] P. W. Langhoff, *J. Phys. Chem.* **100**, 2974 (1996).
- [4] J. A. Sheehy, J. A. Boatz, M. E. Fajardo, and P. W. Langhoff, in *Proceedings of the High Energy Density Matter (HEDM) Contractors' Conference*, P. G. Carrick and N. T. Williams, eds., PL-TR-96-3037 (1997).
- [5] P. W. Langhoff, J. A. Boatz, and J. A. Sheehy, in *Proceedings of the High Energy Density Matter (HEDM) Contractors' Conference*, P. G. Carrick and N. T. Williams, eds., PL-TR-97-3057 (1998).
- [6] F. O. Ellison, *J. Am. Chem. Soc.* **85**, 3540 (1963).
- [7] J. A. Boatz, M. E. Fajardo, and P. W. Langhoff, in *Proceedings of the High Energy Density Matter (HEDM) Contractors' Conference*, P. G. Carrick and N. T. Williams, eds., PL-TR-95-3039 (1996).
- [8] J. A. Boatz, M. E. Fajardo, and J. A. Sheehy, in *Proceedings of the High Energy Density Matter (HEDM) Contractors' Conference*, P. G. Carrick and N. T. Williams, eds., PL-TR-96-3037 (1997).
- [9] J. A. Boatz, J. A. Sheehy, and P. W. Langhoff, in *Proceedings of the High Energy Density Matter (HEDM) Contractors' Conference*, P. G. Carrick and N. T. Williams, eds., PL-TR-97-3057 (1998).
- [10] P. J. Knowles and N. C. Handy, *Chem. Phys. Lett.* **111**, 315 (1984).
- [11] C. E. Morse, *Atomic Energy Levels*, vol. 1 (U. S. Government Printing Office: Washington, DC, 1949).
- [12] R. A. Aziz and M. J. Slaman, *J. Chem. Phys.* **92**, 1030 (1990).
- [13] J. H. Yates, W. C. Ermler, N. W. Winter, P. A. Christiansen, Y. S. Lee, and K. S. Pitzer, *J. Chem. Phys.* **79**, 6145 (1983).
- [14] Gaussian 94, Revision E.2, M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, J. Cioslowski, B. B. Stefanov, A. Nanayakkara, M. Challacombe, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. P. Stewart, M. Head-Gordon, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Pittsburgh PA, 1995.
- [15] R. F. Stewart, *J. Chem. Phys.* **52**, 431 (1970).
- [16] W. Kolos and L. Wolniewicz, *J. Chem. Phys.* **43**, 2429 (1965).

Table I. GTO representations of the radial portions of the hydrogen 1s, 2s, and 2p orbitals. E is the energy of each GTO expansion in atomic units.

Exponent	$\chi_1 = 1s$	$\chi_2 = 2s$	$\chi_3 = 2p$
0.074453	0.19357	0.51052	0.0
0.197572	0.48257	-0.17078	0.0
0.578648	0.33182	-0.13544	0.0
2.071728	0.11354	-0.04207	0.0
11.305637	0.02214	-0.00782	0.0
0.014016	0.0	0.55285	0.16534
0.031833	0.0	0.63915	0.48349
0.076996	0.0	-0.65336	0.36677
0.216081	0.0	-0.11698	0.12355
0.830097	0.0	-0.00513	0.02079
$E =$	-0.499506	-0.124941	-0.124985

Table II. Diffuse GTOs used to augment the hydrogen 1s, 2s, and 2p orbitals listed in Table I.

Function	Symmetry	Contraction coefficients	
		Exponent = 0.02	Exponent = 0.05
ψ_1	s	1.0	0.0
ψ_2	s	0.0	1.0
ψ_3	p	1.0	0.0
ψ_4	p	0.0	1.0
ψ_5	d	0.76530	0.11902

Table III. Diffuse spectral states derived from the GTO expansions in Tables I and II. The energy of each spectral state is given in atomic units.

State	Symmetry	Definition	Energy
χ_4	s	$5.06346 \psi_1 - 2.42337 \psi_2 - 0.54821 \chi_1 - 3.02276 \chi_2$	0.008291
χ_5	p	$1.76217 \psi_3 - 0.72275 \psi_4 - 0.80712 \chi_3$	-0.037286
χ_6	d	ψ_5	-0.051239

Fig. 1. Potential energy of C_{2v} Ar_3 as a function of θ_{123} for three values of γ_0 .

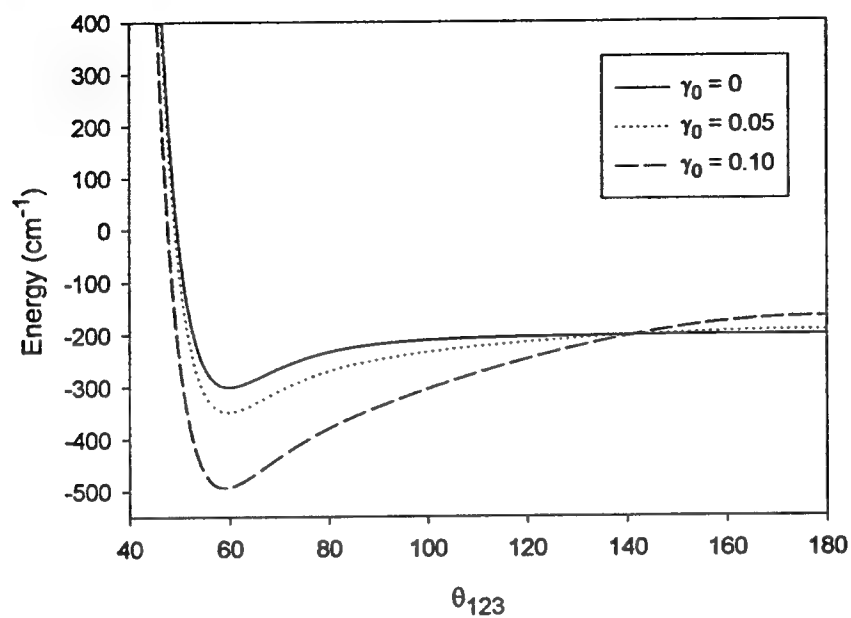


Fig. 2. Potential energy of right-angled Ar_3 as a function of R_{23} for three values of γ_0 .

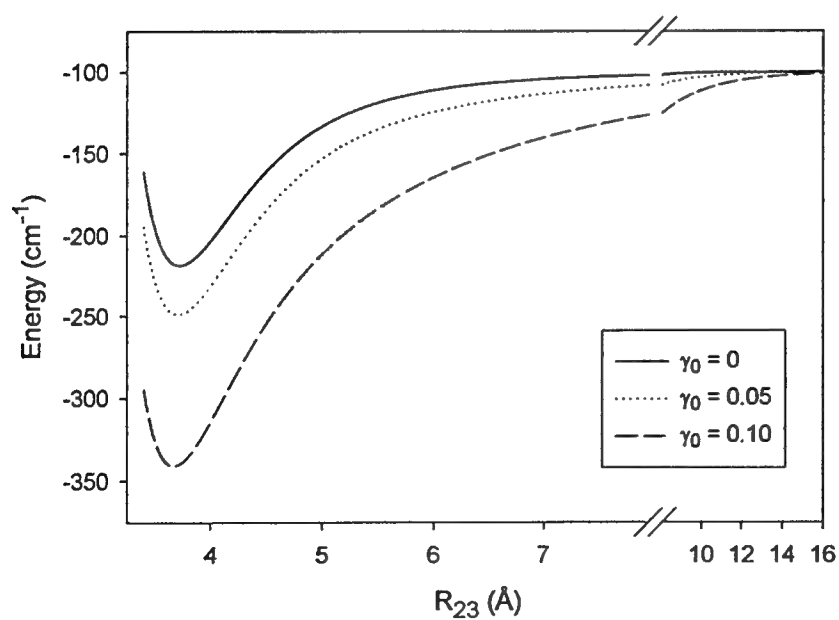


Fig. 3. Potential energy curves for H_2 . The solid lines are the exact $X^1\Sigma_g^+$ and $b^3\Sigma_u^+$ curves. The dotted line is the approximate $X^1\Sigma_g^+$ curve computed with the small spectral product basis set; the dashed line is the approximate $X^1\Sigma_g^+$ curve computed with the large spectral product basis set.

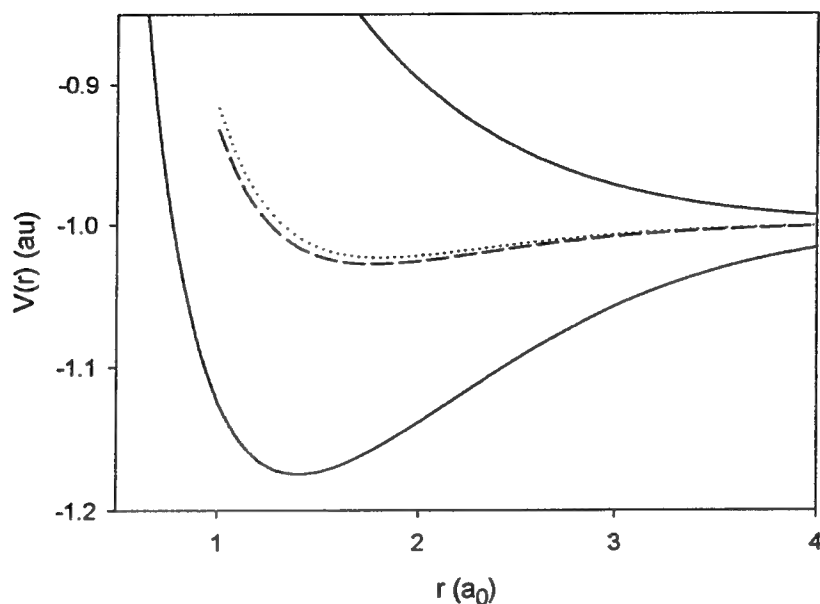
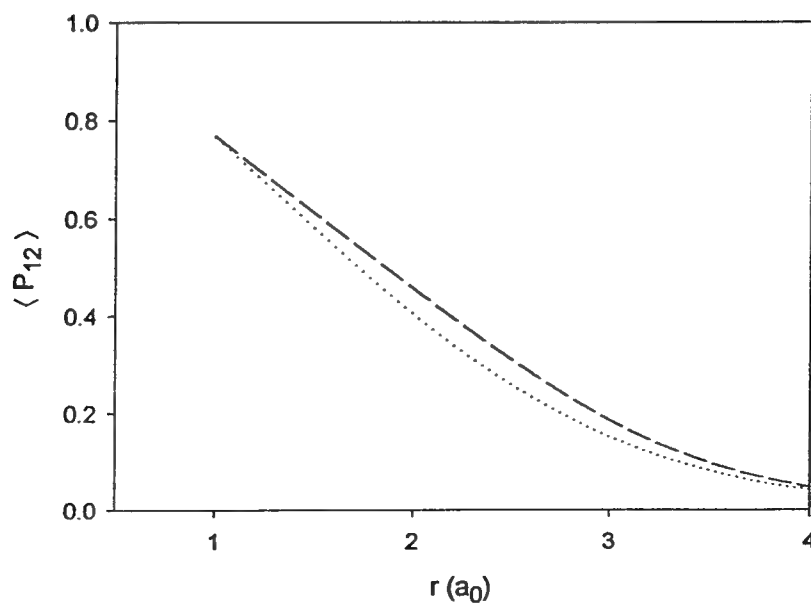


Fig. 4. The expectation value $\langle \hat{P}_{12} \rangle$ for the two spectral product H_2 electronic wavefunctions. The dotted and dashed lines represent the small and large spectral product basis sets, respectively.



**REACTION OF ELECTRONICALLY-EXCITED NITROGEN ATOMS
WITH MOLECULAR OXYGEN**

**Martin A. Hunter
Visiting Assistant Professor
Department of Chemistry**

**Holy Cross College
P.O Box C
1 College Street
Worcester, MA 01610-2393**

**Final Report for:
Summer Faculty Research Program
AFRL/VSBM
Hanscom AFB, MA 01731-3010**

**Sponsored by:
Air Force Office of Scientific Research**

September 1998

REACTION OF ELECTRONICALLY-EXCITED NITROGEN ATOMS WITH MOLECULAR OXYGEN

Martin A. Hunter
Visiting Assistant Professor
Department of Chemistry
Holy Cross College

Abstract

The reaction of O_2 with metastable nitrogen atoms, $N(^2D)$, is a major source of highly-excited NO in the upper atmosphere. We have continued our studies of this reaction at low temperature (80-120 K) in COCHISE, a cryogenic reactor/spectrometer, via the interaction of O_2 and discharge-excited N_2 . Reaction products were monitored by infrared and UV/visible chemiluminescence, indicating significant rovibrational excitation of the $NO(X^2\Pi)$ product. No emissions from $O(^1D)$ product were detected, however, indicating that the major reaction channel ($\phi > 0.7$) involves formation of ground state $O(^3P)$ atoms. Our experiments provide a new estimate of the rate coefficient for this process, $k(100\text{ K}) = (2-5) \times 10^{-11} \text{ cm}^3 \text{ molecule}^{-1} \text{ s}^{-1}$, which is significantly larger than the accepted value at higher temperatures, $k(240-370\text{ K}) = 6 \times 10^{-12} \text{ cm}^3 \text{ molecule}^{-1} \text{ s}^{-1}$. This finding, together with the inference of a near-statistical nascent population distribution of the $NO(X^2\Pi; v, J)$ product, suggest this reaction to proceed via complex-formation without a significant potential energy barrier.

REACTION OF ELECTRONICALLY-EXCITED NITROGEN ATOMS WITH MOLECULAR OXYGEN

Martin A. Hunter

Introduction

The reaction of O_2 with metastable nitrogen atoms, $N(^2D, ^2P)$, plays a significant role in the chemistry of the Earth's upper atmosphere. It is the primary source of thermospheric NO at altitudes above 100 km, and it has also been implicated in the $O(^1D) \rightarrow O(^3P)$ 630 nm emissions observed under both auroral and quiescent conditions in the upper atmosphere.^{1,2} Several aspects of the reaction dynamics of this system, however, remain unresolved. The reactions of both $N(^2D)$ and $N(^2P)$ with O_2 are highly exoergic, and may lead to a variety of electronically excited products:



To date, the relative rates into these various product channels, and their temperature dependence, have not been unambiguously established. Lin and Kaufman first determined the efficient quenching of $N(^2D)$ by O_2 at 300 K to proceed primarily via reaction (1), rather than by simple collisional energy transfer.³ This mechanism is now well established, with a room temperature rate coefficient of $k_1 = 6^{+1}_{-1} \times 10^{-12} \text{ cm}^3 \text{ molecule}^{-1} \text{ s}^{-1}$.^{3,4} However, the relative rates into channels (1a) and (1b) are still unknown. Reaction (2) is even less well characterized. Reported values of its rate coefficient vary over a significant range ($k_2 = (1.8^{+0.2}_{-0.2} - 3.5^{+0.5}_{-0.5}) \times 10^{-12} \text{ cm}^3 \text{ molecule}^{-1} \text{ s}^{-1}$),^{3,5,6} and its various product channels — including the degree of unreactive physical quenching — remain undetermined.

The difficulty in characterizing the kinetics and dynamics of reactions (1) and (2) arises from the

lack of a selective source for either $N(^2D)$ or $N(^2P)$ atoms. Most studies on these systems have relied on mixed $N(^2D)/N(^2P)$ sources, generated by microwave discharge excitation of N_2 /noble gas mixtures. Under these conditions, the metastable N atoms may be individually detected (e.g., via resonance fluorescence,³ resonance-enhanced multiphoton ionization,^{4,5} or spontaneous emission⁶), thus providing reliable measurements for the *total* quenching rates of each of these species by O_2 . However, product state characterization is severely complicated by the combination of mixed reactant source and the availability of multiple reactive channels.

The direct observation of $NO(X^2\Pi)$ product from reactions (1) and (2) was first accomplished by Kennealy and coworkers, who monitored the near-nascent NO vibrational distributions by infrared chemiluminescence.⁷ Their study, based on the interaction between O_2 and discharge-excited N_2 at low temperature (80–120 K), showed the NO product to be highly vibrationally excited. Rate coefficients for the production of $NO(X^2\Pi; v=1-7)$ levels were found to be nearly constant, but decreased monotonically for vibrational levels $v=8-12$. Rawlins *et al.*, using the same apparatus as Kennealy and co-workers, extended their work by further characterizing the NO product rotational distributions.¹ These were observed to be consistently bimodal, with a thermal component around 100 K and a hot contribution at $T_{rot} \sim 10,000$ K. By varying the $[N(^2D)]/[N(^2P)]$ concentration ratio in their reaction chamber, Rawlins *et al.* tentatively assigned the hot rotational distributions to the reaction of O_2 with $N(^2P)$, and the thermal component to reaction with $N(^2D)$. Surprisal analysis of the $NO(X^2\Pi; v, J)$ product state distributions supported this conclusion, and ascribed the marked drop in NO vibrational population above $v=7$ to a large quantum yield ($\Phi_{1b} = 0.76$) for $O(^1D)$ production via reaction (1).

The present work extends the studies of Rawlins *et al.* by *directly* monitoring the production of metastable $O(^1D, ^1S)$ states from reactions (1) and (2). Such a study is of interest in itself, in further clarifying the dynamics of these elementary reaction systems, but it would also have relevance to the aeronomy of the upper atmosphere. Rocket-borne measurements of auroral infrared emissions have confirmed the importance of reaction (1) as a source of highly vibrationally excited $NO(X^2\Pi)$.^{8,9} However, some controversy remains regarding the degree of $O(^1D)$ formation via reaction (1) under thermospheric conditions ($T=300-1500$ K).² A coordinated rocket/satellite measurement of auroral optical emissions by Rees *et al.* first led to the conclusion that the two classical sources of auroral $O(^1D)$ —electron impact excitation of $O(^3P)$ and dissociative recombination of O_2^+ —could not fully account for the observed $O(^1D) \rightarrow O(^3P)$ 630 nm emission rates.^{10,11} Based on the high yield of $O(^1D)$ production inferred from Kennealy *et al.*'s laboratory work, Rusch *et al.* suggested that reaction (1) could account for this "missing source" of 630 nm radiation.¹² This conclusion, however, has been strongly

challenged in light of more recent ground-based and satellite observations of auroral emissions, which place an upper limit of ≤ 0.1 to the quantum yield of $O(^1D)$ from reaction (1) in the auroral thermosphere.²

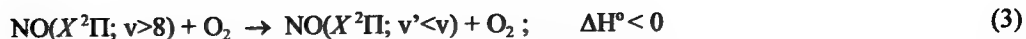
A similar controversy has developed over the role of reaction (1) in the 630 nm airglow observed in the quiescent, unperturbed thermosphere. In their modeling of satellite observations of the dayglow, Torr *et al.* first suggested that reaction (1) was the major source of 630 nm emissions above 150 km.¹³ More recent evidence, however, strongly disputes this claim.^{2,14} Atmospheric Explorer C and E satellite measurements of 630 nm dayglow, for example, show no indication of contributions from reaction (1),¹⁵ while ground-based measurements of 630 nm nightglow at Arecebo have determined reaction (1) to be a minor ($< 5\%$) contributor, even assuming a unity quantum yield into channel (1b).¹⁶

Clearly, a direct measurement of the $O(^3P, ^1D, ^1S)$ quantum yields from reactions (1) and (2) would be instrumental in resolving these questions of aeronomic interest. Lin and Kaufman were the first (and last!) to monitor the $O(^3P)$ product directly, determining a near-unity quantum yield for total O-atom production via reaction (1).³ Their work, however, could not discern the relative contribution of the $O(^1D)$ channel to total O-atom production, due to the efficient quenching of this species by O_2 and N_2 present in their system.

In this paper, we describe the first attempt to directly monitor the production of $O(^1D, ^1S)$ from reactions (1) and (2). Our experiments were conducted in the same cryogenic reaction chamber utilized by Rawlins, Kennealy and coworkers—the COCHISE (cold chemiluminescence infrared stimulation experiment) facility at the Air Force Geophysics Laboratory—¹⁴ adapted for simultaneous collection of both infrared and UV-visible fluorescence. Our results showed no detectable $O(^1D) \rightarrow O(^3P)$ emissions at 630nm, nor $O(^1S) \rightarrow O(^1D)$ emissions at 558nm, placing upper limits of $\Phi_{1b} < 0.3$, $\Phi_{2b} < 0.3$, and $\Phi_{2c} < 0.05$ on the quantum yields of reaction channels (1b), (2b) and (2c), respectively. In addition, we have determined the $NO(X^2\Pi; v, J)$ product state distributions to vary non-linearly with $[O_2]$ concentrations in the reaction chamber: a 100-fold drop in $[O_2]$ reactant concentration leads only to a 50% drop in product $[NO]$, and results in a higher relative population of NO vibrational levels $v=8-12$.

These findings are in conflict with the reaction mechanism previously proposed by Rawlins *et al.*,¹ and we examine an alternative reaction scheme which may account for this discrepancy. Our observations suggest that under conditions of high reactant $[O_2]$ concentration, the nascent $NO(X^2\Pi; v>8)$ populations formed by reactions (1) and (2) may be significantly perturbed by vibrational quenching

with O₂:



If this reaction is assumed to be diffusion-limited (i.e., $k_3 \approx 10^{-10} \text{ cm}^3 \text{ molecule}^{-1} \text{ s}^{-1}$), it could account for the monotonic decrease in NO($X^2\Pi$; $v>7$) vibrational populations observed in our experiments. We describe two processes which may account for such highly efficient vibrational quenching: (i) near-resonant V-V transfer with ground state O₂ molecules, and (ii) transient NO₃ complex-formation.

A detailed kinetic model has been developed to incorporate this vibrational quenching reaction, and to account for the non-linearity in NO(v) production as a function of reactant [O₂] concentration. The latter effect can be shown to result from the interplay between the rate of chemical formation of the NO(v) product and the convective and diffusive losses of these active species away from the spectrometer's field of view. Based on the measured reaction cell pressures and a simple model of the spectrometer's cryopumping dynamics, our kinetic model provides a value of $k(100 \text{ K}) = (2-5) \times 10^{-11} \text{ cm}^3 \text{ molecule}^{-1} \text{ s}^{-1}$ for the rate constant of NO(v) formation. Our model also provides evidence for a near-statistical rotational state distribution of the *nascent* NO(v) product, suggesting that the observed bimodal character of the NO(v, J) distributions arises from efficient rotational quenching of NO(v, J) by the reaction cell bath gases (Ar, O₂ and He) rather than via multiple reaction pathways.

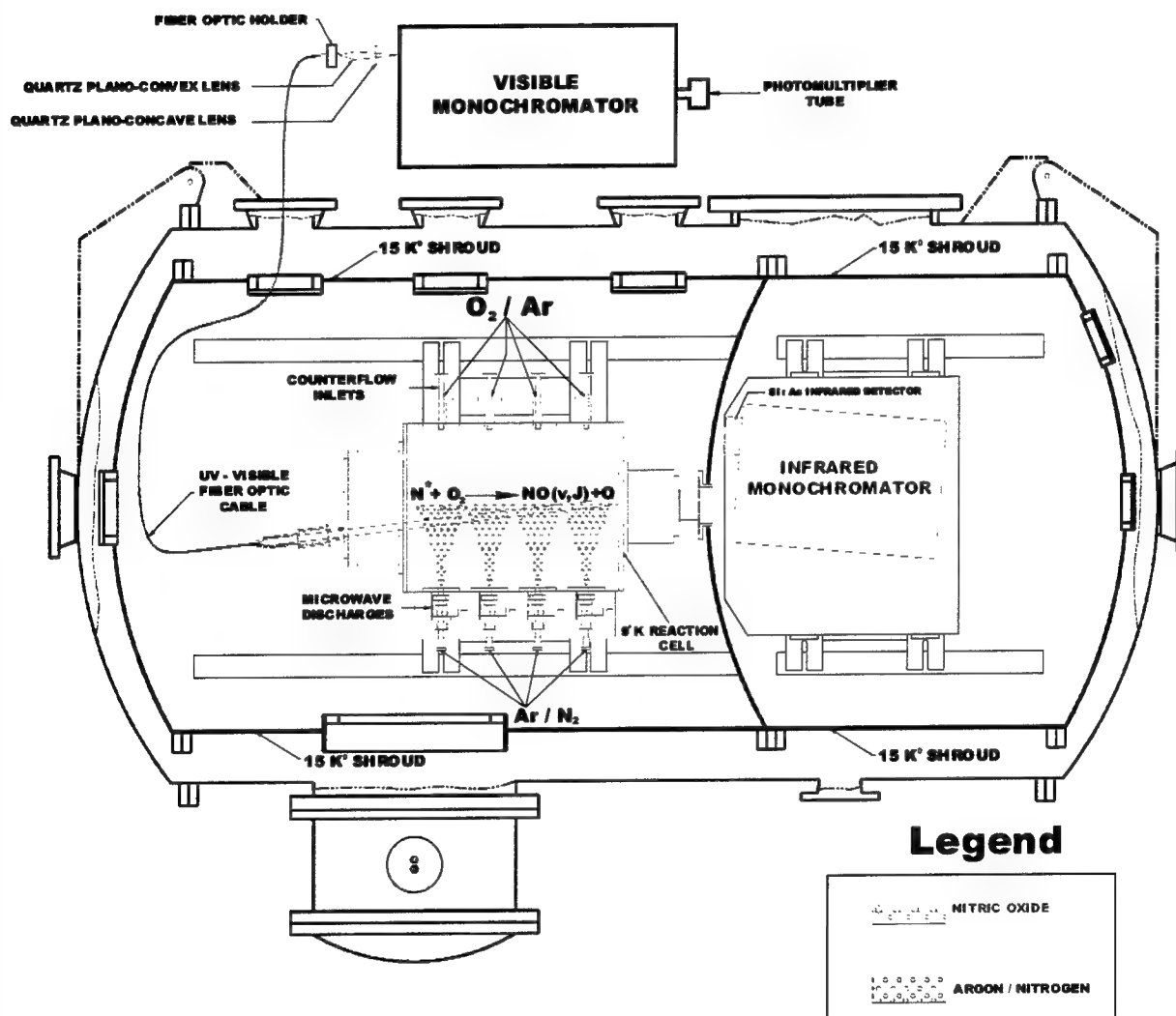
Our most recent experimental findings and kinetic models, therefore, suggest a comprehensive and simplified picture for the reaction dynamics of metastable nitrogen atoms with O₂. In this paper, we demonstrate that our entire set of experimental results can be accounted solely in terms of reaction (1a): the reaction of N(²D) with O₂ to generate NO($X^2\Pi$) and O(³P). The observed bimodal nature of the NO(v, J) rotational distributions in our experiments can be accounted for by efficient rotational quenching of the nascent NO(v, J) product, whereas the sharp monotonic decrease in NO(v) vibrational populations above $v=7$ can be ascribed to efficient vibrational quenching by O₂. There is thus no need to invoke multiple reaction pathways, involving both N(²D) and N(²P), to account for our observed NO(v, J) product state distributions. In addition, our inability to detect any metastable O(¹D, ¹S) emissions places strong restrictions on their production via either reactions (1) or (2). The simplest interpretation of these results must therefore involve reaction (1a) as the principal source of NO in our experiments, since reaction (2a) is disallowed by symmetry restrictions, the rate coefficient of reaction (1) is about 2-3 times faster reaction (2), and the concentration of N(²D) in the reaction cell is expected to be up to an order of magnitude greater than that for N(²P). These findings are in agreement with recent field observations indicating that reaction (1b) is not a significant source of O(¹D) in the upper atmosphere.

Experimental

The experiments were conducted in the same cryogenic reaction chamber utilized by Rawlins, Kennealy and coworkers—the COCHISE (Cold Chemiluminescence Infrared Simulation Experiment) facility at the Air Force Research Laboratory—^{1,4} adapted for simultaneous collection of both infrared (previous configuration) and UV-visible fluorescence. The apparatus was originally designed for nascent-product infrared chemiluminescence spectroscopy studies in a cryogenic environment, under near-single collision conditions, and it has been previously described in detail.¹⁷ A schematic of the COCHISE reaction chamber, including the proposed adaptation to its optical collection system, is shown in Figure 1. The reaction cell is cylindrically symmetric, 60 cm in length and 40 cm in diameter; it is enclosed within a cryogenic thermal shroud and normally maintained at a base temperature of 20 K to eliminate background infrared radiation over the apparatus' operating range of 2-20 μm . The thermal shroud is itself isolated inside a cylindrical main vacuum chamber 3m long and 1.5m in diameter, maintained at 10^{-10} atm by a 50-cm cold-baffled diffusion pump. Reagent gases are introduced into the cryogenic reaction cell through four sets of opposing jets, and interact along the center line of the cell. The techniques of discharge flow and arrested relaxation are combined to allow mixing of the reacting species in the field of view at steady-state pressures around 3-10 mTorr and 100 K temperatures. $\text{N}(^2\text{D})$ and $\text{N}(^2\text{P})$ are produced in our experiments by microwave discharge excitation of 1-10% N_2/Ar gas mixtures at 1 Torr pressures.

In its previous configuration (infrared detection only), a 3"-diameter lens collected radiation along the center line of the reaction cell, focussing the infrared chemiluminescence from a 2.5-liter field of view onto a cryogenic grating monochromator detection system. Our latest addition to this apparatus involves a second set of collection optics for detection capabilities in the visible and UV range. A 1"-diameter lens collects emissions from the reaction zone along an optical axis tilted $\sim 10^\circ$ away from the chamber center line. The light is focussed onto a vacuum-sealed, circular-to-slit silica fiber optic bundle (80 fibers, 110 μm diameter each), and detected externally (i.e., at room temperature) via a 0.50m monochromator through 250 μm slits (0.45nm resolution) with a GaAs photon-counting photomultiplier tube.

Figure 1. Experimental apparatus: COCHISE cryogenic reactor/spectrometer



Results

O(¹D) emits at 630 nm with a radiative lifetime of 148 s, while O(¹S) emits at 556 nm with a lifetime of 800 ms. Based on a calibration of our collection efficiencies using a 200 W standard NIST lamp, our detection limits for O(¹D) and O(¹S) in our experiments were determined to be 10^6 and 10^8 cm⁻³, respectively, per minute of integration time. Under typical COCHISE operating conditions, the total density of NO(v) product is on the order of 10^9 cm⁻³. Thus, if Rawlins' prediction of a quantum yield of 0.76 for reaction (1b) were correct,¹ we should have observed 630 nm emissions with a signal-to-noise ratio of about 8:1. The lack of any such signal places an upper limit to the quantum yield for O(¹D) generation via reactions (1) and (2) to <0.3, even after accounting for the efficient secondary quenching of O(¹D) with O₂ reactant. A similar null result for 556 nm emissions places an upper limit of <0.05 for the quantum yield for O(¹S) generation via reactions (1) and (2). These results are significantly in disaccord with the kinetic scheme presented by Rawlins *et al.* for reactions (1) and (2).¹

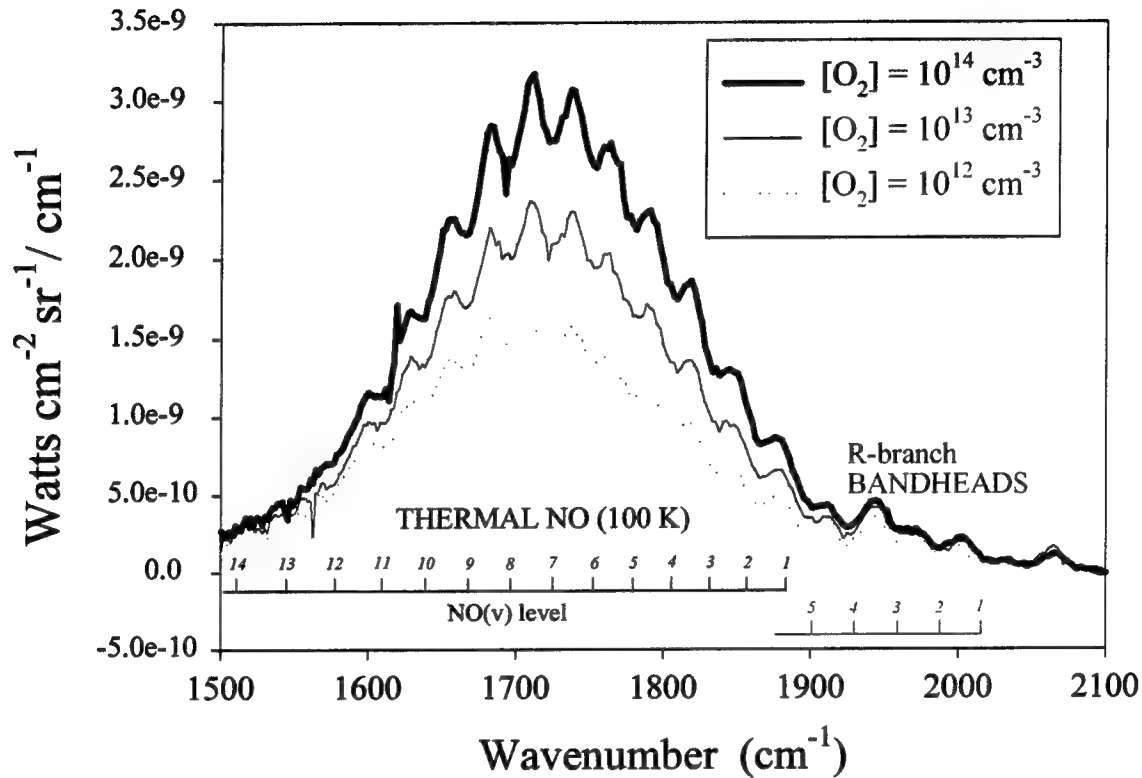
Improvements in our apparatus' operating range have also allowed us to collect NO(v) infrared emission spectra over a broader range of O₂ reactant concentrations than those reported by Rawlins and co-workers.¹ Our main results are depicted in Figures 2 and 3, which show NO(v) fundamental emission spectra as a function of [O₂] in the reaction chamber. The most striking features of these new results are:

- (i) NO(v) production shows a strongly non-linear dependence on [O₂]: a 100-fold drop in [O₂] results in only a 50% drop in NO(v) emissions.
- (ii) The observed NO(v) populations vary as a function of [O₂], showing a relative increase in production of NO(v>7) vibrational levels as the [O₂] reactant density is reduced.

Neither of these effects can be accounted for by the kinetic scheme for reactions (1) and (2) proposed by Rawlins and co-workers.

Process (i) may be understood by considering the interplay between the rate of chemical formation of the NO(v) product and its removal from the spectrometer's field of view via diffusive and convective losses. If the loss mechanisms are relatively slow, leading to long residence times in the reactor's interaction region, it may become possible for a fast reaction to achieve steady state conditions before complete removal of reactants from the field of view. In this case, there will no longer be a linear relationship between reactant concentrations and observed product emissions under pseudo-first order conditions. The complex nature of the cryogenic environment in the COCHISE reactor precludes an exact calculation of the residence times of active species in our chamber's field of view. However, based on our measured cell pressures of 3-10 mtorr and the total flow rate of gas into the reactor (8 standard

Figure 2. NO(v) emission spectra as a function of [O₂]: Experimental results



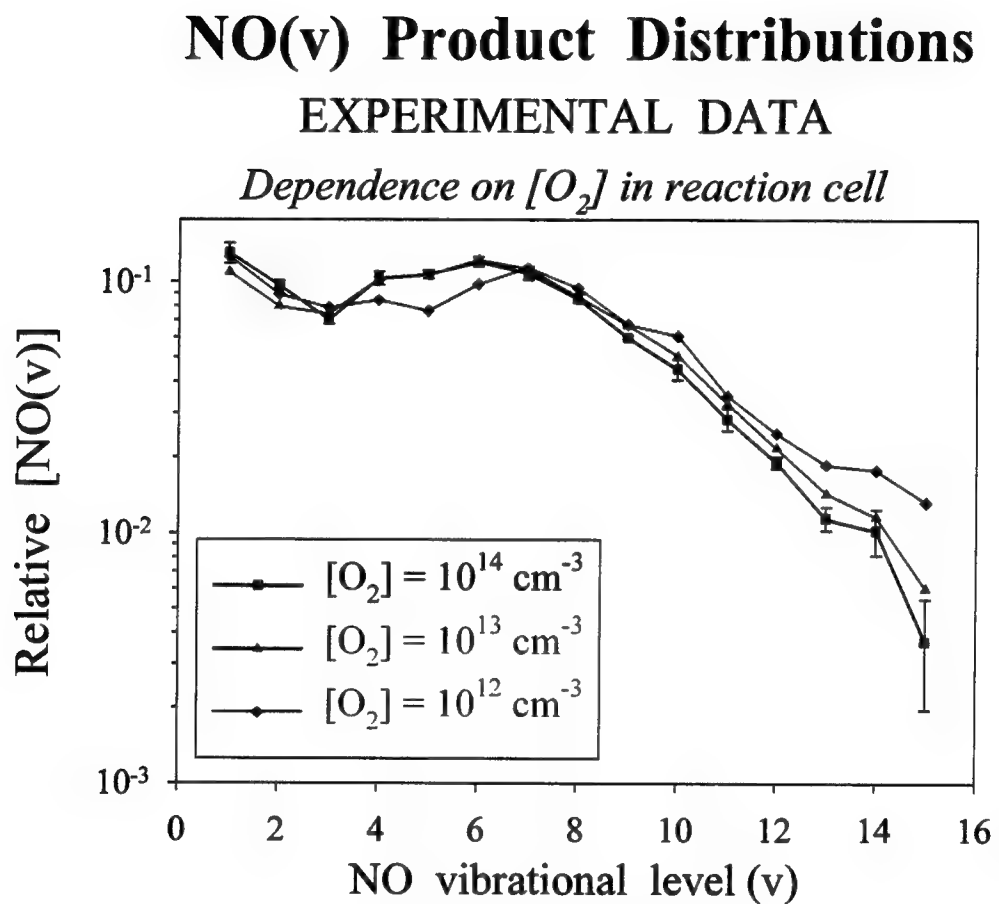
- Total (integrated) NO(v) emission exhibits strongly NON-LINEAR dependence on [O₂]

$$\tau_{\text{formation}}(N^* + O_2) < \tau_{\text{loss}}(\text{diffusion})$$

- Relative NO(v) population distribution varies with [O₂]

$$\uparrow [O_2] \Rightarrow \downarrow \text{NO vibrational excitation}$$

Figure 3. NO(v) vibrational product state distributions as a function of $[O_2]$



liters per minute), it is possible to estimate the residence time to 1-2 ms within the 2.5-l interaction region.¹⁸ Under these conditions, our experimental results can be quantitatively modeled by assuming that the rate constant for NO(v) formation (via either reaction (1) and/or (2)) is $k(100\text{ K}) = (2-5) \times 10^{-11} \text{ cm}^3 \text{ molecule}^{-1} \text{ s}^{-1}$. This value is considerably greater than the accepted rate coefficient for these reactions at room temperature. Despite the large uncertainty in our measurement, we are confident that the difference is significant. Before reporting this value to the academic community via a formal publication, however, it will be necessary to better characterize the cell pressures in our reactor (and thereby reduce the uncertainty in our estimated residence times in the interaction region). We are currently working toward this goal; it is a complex task due to the cryogenic environment of our apparatus, as well as the non-static nature of the gas in our cryo-pumped system. We note that recent experimental and theoretical work by other groups, however, supports the tendency for reaction (1) to increase in rate as the temperature is lowered, although to a lesser extent than observed in our system.¹⁹

Our observed dependence of the NO(v) vibrational distributions on $[\text{O}_2]$ (Figure 3) clearly indicates significant quenching of NO(v>7) by molecular O_2 under our experimental conditions. In order to quantitatively account for the observed degree of quenching, however, it is necessary for reaction (3) to proceed at near gas-kinetic rates. Previous measurement of this rate coefficient, for NO(v) levels $v < 8$, have determined this reaction to be slow, proceeding with rate coefficients $k < 2 \times 10^{-12} \text{ cm}^3 \text{ molecule}^{-1} \text{ s}^{-1}$ at 300 K. However, two mechanisms exist which may lead to an enhanced vibrational quenching rate for NO(v) levels above $v=8$. Resonant V-V transfer with ground state O_2 molecules can occur with a quantum defect of only 14 cm^{-1} for $v=12$ (increasing to 97 cm^{-1} by $v=9$), and could thus easily account for the monotonic decrease in our observed NO(v) populations above $v=7$. In addition, NO(v) vibrational levels above $v=8$ may surmount the activation barrier for NO_3 complex-formation, and may thus provide an additional mechanism for enhanced vibrational relaxation upon collisions with molecular O_2 .

A summary of these results, and the kinetic scheme utilized to interpret them, is given in Figure 4.

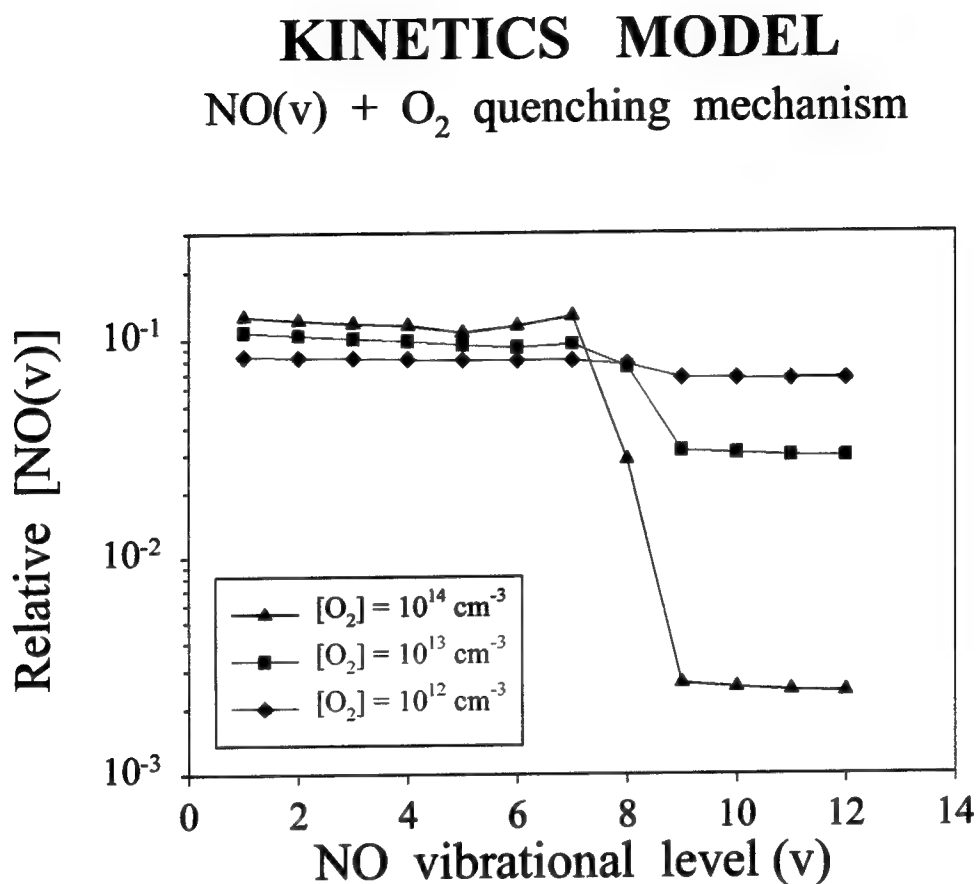
Finally, we note that we have also been developing a detailed kinetic scheme to account for the rotational, as well as the vibrational, NO product state distributions in our experiments. Our results are only preliminary, and the model shall be further refined in the future. However, to a first approximation, it appears that the bimodal NO(v,J) rotational distributions traditionally observed in COCHISE experiments can be explained in terms of efficient rotational quenching of a broadly-populated (possibly

statistical) nascent NO(v,J) distribution. Future work will thus establish the possibility of a single reaction mechanism to account for all of our experimental results in COCHISE i.e., metastable N atoms reacting with O₂ along a single potential energy surface, to yield a near-statistical nascent NO(v,J) state distribution which is significantly affected by both rotational and vibrational quenching effects under our experimental conditions. Considerations of expected reactant densities in COCHISE,¹ as well as conservation of symmetry effects, strongly suggest this single source of NO(v,J) must involve the reaction of O₂ with N(²D), rather than with N(²P). Since no metastable O atom emissions were observed in our experiments, this would suggest that reaction (1a) is the principal reaction mechanism for NO(v,J) formation at 100 K. Its relatively fast rate coefficient, and near-statistical product energy disposal, would strongly suggest the reaction to proceed via NO₂ complex-formation without a significant potential barrier in the entrance channel.

Conclusion

The reaction of O₂ with metastable nitrogen atoms, N(²D) and N(²P), was studied at low temperature (80-120 K) in COCHISE, a cryogenic reactor/spectrometer, via the interaction of O₂ and discharge-excited N₂. Reaction products were monitored by infrared and UV/visible chemiluminescence, indicating significant rovibrational excitation of the NO(X²Π) product. No emissions from O(¹D) product were detected, however, indicating that the major reaction channel ($\Phi > 0.7$) involves formation of ground state O(³P) atoms. Our experiments provide a new estimate of the rate coefficient for this process, $k(100\text{ K}) = (2-5) \times 10^{-11} \text{ cm}^3 \text{ molecule}^{-1} \text{ s}^{-1}$, which is significantly larger than the accepted value at higher temperatures, $k(240-370\text{ K}) = 6 \times 10^{-12} \text{ cm}^3 \text{ molecule}^{-1} \text{ s}^{-1}$. This finding, together with the inference of a near-statistical nascent population distribution of the NO(X²Π;v,J) product, suggest this reaction to proceed via complex-formation without a significant potential energy barrier. These results are in qualitative agreement with recent field measurements of the dayglow and auroral events, which have placed significant restrictions on the possibility of O(¹D) generation via the reaction of O₂ with metastable nitrogen atoms in the upper atmosphere. We have also found evidence for efficient (near gas-kinetic) vibrational quenching of NO(v>8) with ground state O₂, and are actively pursuing the possibility of efficient rotational quenching of NO(v=1-8; J<100) by collisions with O₂, He and Ar.

Figure 4. Kinetic scheme for modeling NO(v) populations as a function of [O₂] in COCHISE



- (1) $N^* + O_2 \rightarrow NO(v) + O$; $k_1 = 5 \times 10^{-11} \text{ cm}^3 \text{ molecule}^{-1} \text{ s}^{-1}$
- (2) $NO(v > 8) + O_2 \rightarrow NO(v' < v) + O_2$; $k_2 = 10^{-10} \text{ cm}^3 \text{ molecule}^{-1} \text{ s}^{-1}$
- (3) $NO(v \leq 8) + O_2 \rightarrow NO(v' < v) + O_2$; $k_3 < 10^{-12} \text{ cm}^3 \text{ molecule}^{-1} \text{ s}^{-1}$

$$[N^*] = 10^{10} \text{ cm}^{-3}$$

$$P_{\text{cell}} = 10 \text{ mtorr} \quad (\tau_{\text{diffusion}} \approx 2.5 \text{ ms})$$

Nascent NO(v) distribution: v-independent

Acknowledgments

I am grateful for the substantial help and support provided by my co-workers at the AFRL/VSBM facility: Dr. S. M. Miller, Dr. S. J. Lipson, Dr. P. S. Armstrong, Dr. R. Lockwood, Dr. J. A. Dodd, J. Cappelli and V. Kennedy. Financial support was provided by the AFOSR through a Summer Faculty Research Grant, and was coordinated via Research & Development Laboratories (RDL).

- ¹W. T. Rawlins, M. E. Fraser and S. M. Miller, "Rovibrational excitation of nitric oxide in the reaction of O₂ with metastable atomic nitrogen", *J. Phys. Chem.*, **93**, 1097 (1989), and references quoted therein.
- ²R. Link and P. K. Swaminathan, "N(²D) + O₂: A source of thermospheric 6300 Å emission?", *Planet. Space Sci.*, **40**, 699 (1992), and references quoted therein.
- ³C. L. Lin and F. Kaufman, "Reactions of metastable nitrogen atoms", *J. Chem. Phys.*, **55**, 3760 (1971).
- ⁴C. Fell, J. I. Steinfeld and S. Miller, "Quenching of N(²D) by O(³P)", *J. Chem. Phys.*, **92**, 4768 (1990), and references therein.
- ⁵C. M. Phillips, J. I. Steinfeld and S. M. Miller, "Remeasurement of N(²P) + O₂ reaction rate using multiphoton ionization detection of nitrogen atoms", *J. Phys. Chem.*, **91**, 5001 (1987).
- ⁶L. G. Piper, "The reactions of N(²P) with O₂ and O", *J. Chem. Phys.*, **98**, 8560 (1993), and references therein.
- ⁷J. P. Kennealy, F. P. Del Greco, G. E. Caledonia and B. D. Green, "Nitric oxide chemiexcitation occurring in the reaction between metastable nitrogen atoms and oxygen molecules", *J. Chem. Phys.*, **69**, 1574 (1978).
- ⁸W. T. Rawlins, G. E. Caledonia, J. J. Gibson and A. T. Stair, Jr., "Infrared emission from NO($\Delta v=1$) in an aurora: spectral analysis and kinetic interpretation of HIRIS measurements", *J. Geophys. Res.*, **86**, 1313 (1981).
- ⁹G. E. Caledonia and J. P. Kennealy, "NO infrared radiation in the upper atmosphere", *Planet. Space Sci.*, **30**, 1043 (1982).
- ¹⁰M. H. Rees, A. I. Stewart, W. E. Sharpe, P. B. Hays, R. A. Hoffman, L. H. Brace, J. P. Doering and W. K. Peterson, "Coordinated rocket and satellite measurements of an auroral event. 1. Satellite observation and analysis", *J. Geophys. Res.*, **82**, 2250 (1977).
- ¹¹W. E. Sharpe, M. H. Rees and A. I. Stewart, "Coordinated rocket and satellite measurements of an auroral event. 2. The rocket observation and analysis", *J. Geophys. Res.*, **84**, 1977 (1979).
- ¹²D. W. Rusch, J. -C. Gérard and W. E. Sharpe, "The reaction of N(²D) with O₂ as a source of O(¹D) atoms in aurorae", *Geophys. Res. Lett.*, **5**, 1043 (1978).
- ¹³D. G. Torr, P. G. Richards and M. R. Torr, "Destruction of N(²D) by O₂: a major source of 6300-Å dayglow emission", *Geophys. Res. Lett.*, **7**, 410 (1980).
- ¹⁴R. Link, J. C. McConnell and G. G. Shepherd, "A self-consistent evaluation of the rate constants for the production of the 6300 Å airglow", *Planet. Space Sci.*, **29**, 589 (1981).
- ¹⁵S. C. Solomon and V. J. Abreu, "The 630 nm dayglow", *J. Geophys. Res.*, **94**, 6817 (1988).
- ¹⁶R. Link and L. L. Cogger, "A reexamination of the O I 6300-Å nightglow", *J. Geophys. Res.*, **93**, 9883 (1988).
- ¹⁷W. T. Rawlins, H. C. Murphy, G. E. Caledonia, J. P. Kennealy, F. X. Robert, A. Corman and R. A. Armstrong, "COCHISE: laboratory studies of atmospheric IR chemiluminescence in a cryogenic environment", *Applied Optics*, **23**, 3316 (1984).
- ¹⁸G. E. Caledonia, B. D. Green, G. A. Simons, J. P. Kennealy, F. X. Robert, A. Corman and F. P. Delgreco, "COCHISE Studies I: Fluid Dynamical and Infrared Spectral Analysis", AFGL-TR-77-0281, Environmental Research Papers, No. 619 (1977).

- ¹⁹Y. Shihira, T. Suzuki, S. Unayama, H. Umemoto and S. Tsunashima, "Reactions of N(2^2D) and N(2^2P) with O₂", *J. Chem. Soc. Faraday Transactions*, 90(4), 549 (1994).
- ²⁰B. D. Green, G. E. Caledonia, R. E. Murphy and F. X. Robert, "The vibrational relaxation of NO($v=1-7$) by O₂", *J. Chem. Phys.*, 76(5), 2441 (1982).

DETERMINISTIC METHODS FOR BLIND
RESTORATION OF ADAPTIVE OPTICS IMAGES
OF SPACE OBJECTS

Brian D. Jeffs
Associate Professor
Department of Electrical and Computer Engineering

Brigham Young University
459 CB, Provo, UT 84602, USA.
E-mail: bjeffs@ee.byu.edu

Final Report for:
Summer Faculty Research Program
Air Force Research Laboratory, AFRL/DES

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

AFRL

September 1998

DETERMINISTIC METHODS FOR BLIND RESTORATION OF ADAPTIVE OPTICS IMAGES OF SPACE OBJECTS

Brian D. Jeffs

Associate Professor

Department of Electrical and Computer Engineering
Brigham Young University

Abstract

The focus of this research program is algorithm development for blind image restoration of multiframe adaptive optics (AO) telescope images. In particular, methods of primary interest are those algorithms capable of removing residual blur in space objects images acquired using the Starfire Optical Range ground based 1.5m and 3.5 m AO telescope systems located at Kirtland Air Force Base, NM. Deterministic, or algebraic methods which can sometimes lead to closed form solutions rather than iterative algorithms are emphasized in this research. New techniques drawn from the literature on blind channel equalization for digital communications were evaluated for adaptation to the blind restoration problem. Three methods were developed to the simulation stage: 1) Homomorphic blind star image restoration, which uses an extremely fast closed form computation to recover blurred points, 2) Direct multiframe solution, which finds the linear algebraic nullspace of the observed data frames, from which the true image is found by subspace matching, and 3) An algebraic cross relationship algorithm which enables estimating the unknown blurs without having to solve for the object image. Adaptation of this method to a previous iterative Bayesian algorithm was also investigated. Ongoing research activities are described.

DETERMINISTIC METHODS FOR BLIND RESTORATION OF ADAPTIVE OPTICS IMAGES OF SPACE OBJECTS¹

Brian D. Jeffs

1 Introduction

1.1 Background

This report discusses progress made during the 1998 AFOSR-SFRP tour of Dr. Brian D. Jeffs while at the Starfire Optical Range, Air Force Research Laboratories, Kirtland Air Force Base, New Mexico. The focus of this research program is algorithm development for blind image restoration of multiframe adaptive optics (AO) telescope images. In particular, methods of primary interest are those algorithms capable of removing residual blur in space objects images acquired using the SOR ground based 1.5m and 3.5 m AO telescope systems.

This work is a continuation of studies undertaken during a 1997 AFOSR-SFRP tour, where statistical estimation algorithms incorporating image prior models for object and blur were developed. The emphasis of this year's research has shifted to considering deterministic, or algebraic methods which can sometimes lead to closed form solutions rather than iterative algorithms. Many such techniques have been introduced over the past few years in the literature on blind channel equalization for digital communications [1, 2, 3, 4, 5, 6, 7, 8, 9]. A major focus of this present study is to evaluate how effectively these algorithms can be adapted to the blind AO multiframe image restoration problem.

AO systems can remove much of the atmospheric-turbulence-induced blur in astronomical and orbiting space objects imagery, but there is typically some adaptation error [10] [11] [12] [13]. Thus, a residual time-varying blur below the theoretical aperture cutoff resolution limit remains. Though the general structure of this residual blur is known [14] [15], the detail is not, and varies significantly over a period of milliseconds.

It is often impractical to acquire an accurate sample of the true residual psf by observing a nearby star. This residual blur is spatially varying, so a natural guide star may not be in the same isoplanatic patch as the object, and in the case of tracking space objects, no natural reference stars which follow the object are available. When observing bright objects, it is possible to acquire a sequence of high SNR, closely spaced (in time) frames with significant blur variation from frame to frame. This scenario is ideally suited to multiframe blind deconvolution methods which exploit the blur diversity to solve the difficult blind problem.

As is usually the case, the blind restoration problem addressed here is complicated by the ill-posed nature of the associated inverse problem [16], and by the fact that blur psf, H , and the true object image, F can only be identified in combination ($H * F$) unless distinguishing prior information is available. This solution ambiguity is due to the commutative property of convolution which will yield the same observation regardless of whether a given image feature is contained in H or F . Therefore, the blind restoration problem is hopeless without a well formed method of expressing and imposing prior information on the solution. A number of authors have successfully applied various "hard" and "soft" constraints on the object and blur solutions as a means of incorporating prior information into the blind multiframe solution [11] [12] [10] [13] [17] [18] [19] [15].

The availability of multiple observations with different but unknown blur psf's, H_i , can be a great aid in estimating F . The multiple frame observation reduces convolutional ambiguity because any image component that is not common to all frames cannot be a component of the F . The method presented below will be developed to exploit this fact, and will be applicable for both single and multiple frame cases.

¹Sponsored by the Air Force Office of Scientific Research, Bolling Air Force Base, DC. The author acknowledges the contributions to this work from Dr. Julian C. Christou and Dr. Robert Fugate, Starfire Optical Range, Air Force Research Laboratory.

1.2 Problem Statement

We adopt the following familiar linear image observation model

$$G_i = H_i * F + \eta_i, \quad i = 1 \dots M \quad (1)$$

where G_i is the i^{th} observed image frame, η_i is additive noise, and $*$ denotes 2-D convolution. G_i , H_i , F , and η_i are all assumed to be realizations of 2-D random fields, but F remains constant between observation frames. It will also be convenient in our analysis to use the vector-matrix forms of equation (1). In this notation G_i , H_i , and F may be in column scanned vector form (\mathbf{g}_i , \mathbf{h}_i and \mathbf{f} respectively), or represented as large Toeplitz-Block-Toeplitz convolution matrices (\mathbf{G}_i , \mathbf{H}_i and \mathbf{F}) [20]. For example, equation (1) can be written as

$$\mathbf{g}_i - \mathbf{H}_i \mathbf{f} = \eta_i, \quad i = 1 \dots M, \quad \text{or} \quad (2)$$

$$\mathbf{g}_i - \mathbf{F} \mathbf{h}_i = \eta_i, \quad i = 1 \dots M \quad (3)$$

When it is necessary to represent the entire set of observed image frames as a single entity, equation (2) is represented with the following notation

$$\begin{aligned} \bar{\mathbf{g}} &= \mathcal{H} \mathbf{f} + \bar{\eta}, \quad \text{where} \\ \bar{\mathbf{g}} &= [\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_M]^T, \quad \bar{\eta} = [\eta_1, \eta_2, \dots, \eta_M]^T, \\ \mathcal{H} &= \begin{bmatrix} \mathbf{H}_1 \\ \vdots \\ \mathbf{H}_M \end{bmatrix}. \end{aligned} \quad (4)$$

With the mathematical notation established, we are now ready to express a statement of the problem: "Given observations $\bar{\mathbf{g}}$, with \mathcal{H} unknown, estimate \mathbf{f} ."

The following sections describe progress in several different, yet related areas of research for blind multiframe image restoration. Each of these topics was studied during the 1998 SFRP tour and has shown promise for future investigation. My emphasis during this tour was to investigate new, unexploited technologies and algorithms for blind AO image restoration. In particular, deterministic (non statistical estimation theory) methods drawn from blind channel estimation for digital communications were considered to determine if they could be adapted to the AO restoration problem. Sufficient progress was made in the areas of homomorphic blind star image restoration, direct multiframe object solutions, and cross relationship blur estimation to warrant continued study.

2 Homomorphic Blind Star Image Restoration

This method of deterministic blind restoration is based on filtering the homomorphic transform known as the 2-D complex cepstrum [21, 22]. This approach relies on a property of the cepstrum that in the transform domain, convolutional terms are mapped to additive terms. For example, as shown in equation (5), the cepstral transforms, \hat{H}_i and \hat{F} are added to form \hat{G}_i , rather than convolved as in equation (1).

$$\begin{aligned} G_i &= H_i * F \\ \hat{G}_i &\equiv \mathcal{F}^{-1} \{ \log \{ \mathcal{F} \{ G_i \} \} \} \\ &= \hat{H}_i + \hat{F} \end{aligned} \quad (5)$$

where $\mathcal{F}\{\}$ indicates the Fourier transform. If \hat{H}_i and \hat{F} have non-overlapping regions of support, a simple masking operation can be used to select the desired components, and form an estimate of F using the inverse cepstrum transform as in equation (6). W is a masking window with entries of "1" or "0" used to select the desired components.

$$\bar{F} = \mathcal{F}^{-1} \{ \exp \{ \mathcal{F} \{ W \cdot \hat{G}_i \} \} \} \quad (6)$$

Clearly this method will only work if \hat{H}_i and \hat{F} are essentially non-overlapping in the transform index, quefrequency. In the case of point-like star images, we have found that \hat{F} dominates the high "quefrequency" terms, while \hat{H}_i , which

is low-pass in nature, resides in the low "quefrequency" region. Thus W need only block the low terms to form an estimate of F . This method is related to techniques used in pitch period estimation for speech coding [21].

Figure 1 illustrates application of this technique to star image restoration. A single synthetic star image was blurred and noise corrupted to simulate an AO frame. The processing of equations (5) and (6) were then applied to restore the image. Note that in the result, most of the original stars were recovered with high resolution, but that additional noisy "star" artifacts also appear.

Advantages of this method are principally related to its simplicity. The closed form solution does not require iteration, and using FFT code for the transforms makes it extremely efficient. A 256 by 256 pixel image can be processed in seconds on a modest p.c. Additionally, the resultant image is high resolution, with no blurring components.

The biggest disadvantages of the approach is its sensitivity to noise. As seen in Figure 1, noise produces false star artifacts. Also, there is at present no obvious method for jointly processing multiple frames, each with distinct blur, in order to reduce noise effects. Simply averaging a series of restored frames does reduce noise, but blurs the stars due to jitter in the estimated star positions. Future work will attempt to generalize the homomorphic method for multiframe data, and to reduce artifacts. With the existing limitations, this method appears to be primarily useful as a fast computation initialization for other iterative blind restoration algorithms.

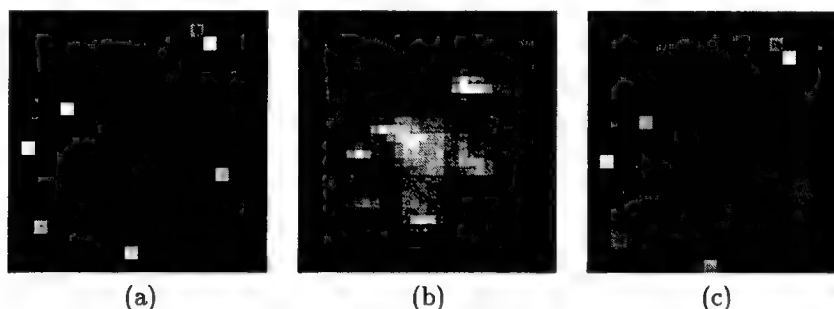


Figure 1: A Single frame example of homomorphic blind star image restoration. a) The original (synthetic) star image. b) Blurred and noise corrupted image. c) The restored image.

3 Direct Multiframe Solution

In recent years there have been a number of articles in the digital communication literature that propose deterministic and subspace decomposition methods for solving the blind channel equalization problem [1, 2, 3, 4, 5, 6]. For a typical scenario, there are multiple sensors which receive a desired coded signal over radio frequency (r.f.) channels with unknown propagation distortion. These channels are represented for each sensor as FIR convolutional filters with distinct, impulse responses corresponding to the unknown r.f. multipath distortion. The objective is to equalize, or remove the effect of channel distortions to enable accurate decoding of the desired signal. This problem is structurally similar to multiframe blind image restoration. The true image, f , corresponds to the coded signal, the observed image frames, g_i , correspond to the receiver sensor signals, and the unknown blurs, H_i , corresponds to the unknown radio frequency channel responses.

The potential advantage of these new methods is that they may yield more accurate solutions than existing image restoration algorithms, and may do so using closed form solutions rather than the iterative, slowly convergent optimization algorithms typically used. Additionally, these methods are direct estimators of f , and do not require a joint estimate of blurs H_i . This drastically reduces the degrees of freedom in the estimation problem, and therefore may reduce estimator error.

The focus of this present research topic is to generalize and exploit some of these channel equalization methods for use in multiframe blind image restoration. A few authors have proposed adaptations of blind channel equalization for image restoration [9, 23]. It is our opinion that these methods, as presented, have computational resource requirements that make their use for AO image restoration impractical. The Emphasis of our research will be to exploit underlying principles of blind equalization methods in a more computationally tractable form. This may involve a hybrid of iterative optimization and closed form solution techniques. Also, in the case of point source

images like star fields, it is expected that the underlying special data structure may be exploited to reduce memory size and computational load.

It has been shown that if the \mathbf{H}_i have no common factors, and the signal, \mathbf{f} has sufficient modes, then \mathbf{f} can be uniquely determined in the absence of noise with no knowledge of \mathbf{H}_i [8, 1]. Typically, this translates to requiring that three or more frames with sufficiently distinct blurs be available. As an initial study of this property, we have used the method of Liu and Xu for blind recovery of point source images.

The Liu - Xu method begins by forming a block Hankel matrix from the observed images

$$\mathbf{G}_{(K)} = \begin{bmatrix} \bar{\mathbf{g}}_1 & \cdots & \bar{\mathbf{g}}_{N-K+1} \\ \vdots & & \vdots \\ \bar{\mathbf{g}}_K & \cdots & \bar{\mathbf{g}}_N \end{bmatrix}, \text{ where} \quad (7)$$

$$\bar{\mathbf{g}}_k = [\mathbf{g}_1(k), \cdots, \mathbf{g}_M(k)]^T.$$

Here index k selects a particular pixel across all M frames, N is the total number of pixels in G_i , L is the number of pixels in H_i , and K is the smoothing factor chosen by the analyst. A singular value decomposition is then used to compute \mathbf{V}_o , the nullspace matrix for $\mathbf{G}_{(K)}$. The nullspace matrix for \mathbf{f} is then constructed as

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_o & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_o & \mathbf{0} & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{V}_o \end{bmatrix} \quad (8)$$

where $\mathbf{0}$ is an $(N - 2L - 2K - 1) \times 1$ vector. \mathbf{f} is then found simply as the non-trivial solution to $\mathbf{V}\mathbf{f} = \mathbf{0}$.

Figures 2 and 3 illustrate how this algorithm works for a 1-D point source (star) image. Figure 2a shows plots for the six frames of observed data, while 2b shows the six corresponding PSF's. Note that each PSF is a Gaussian function with randomly selected variance (width) and mean (shift position) values. Figure 3a shows the true image, with four distinct point (star) sources, and 3b shows the reconstructed \mathbf{f} computed from the data of Figure 2a alone, with no knowledge of the blur PSF's. All four of the point sources were correctly located.

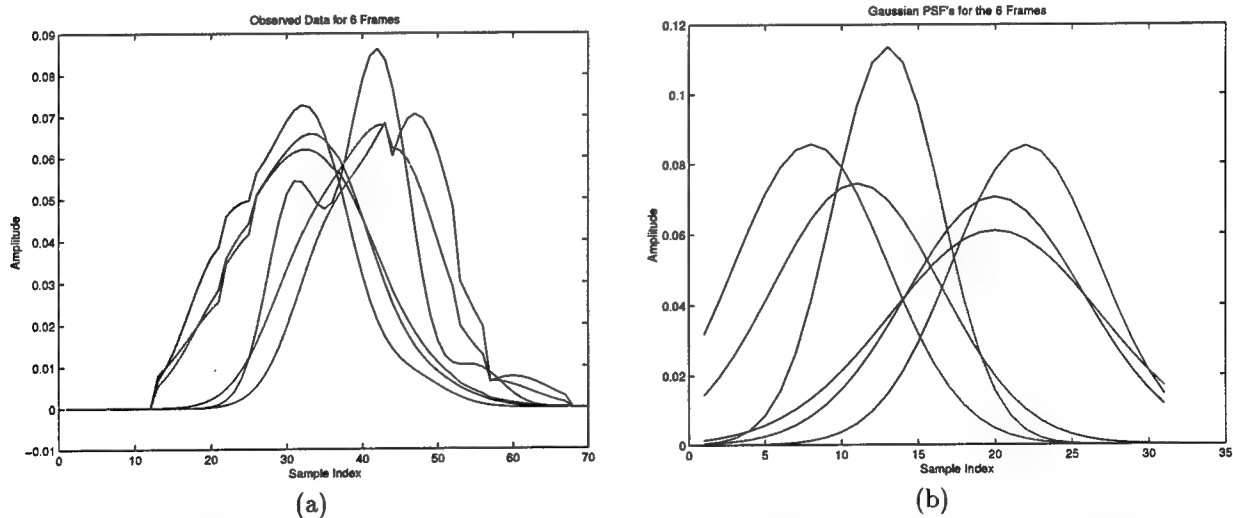


Figure 2: An example of the direct solution method for 1-D multiframe point source image restoration a) The observed data for each of six frames (over-plotted). PSNR is 200 dB. b) Blur point spread functions for the six frames (over-plotted).

It is noteworthy in the example shown that the solution is obtained with a completely closed form computation, and without any estimate of the blur PSF's being formed. These initial tests were restricted to 1-D examples due to the large size of $\mathbf{G}_{(k)}$ used in eigenvector computation. In future work we will exploit the prior knowledge that the true image is point-like to simplify computation.

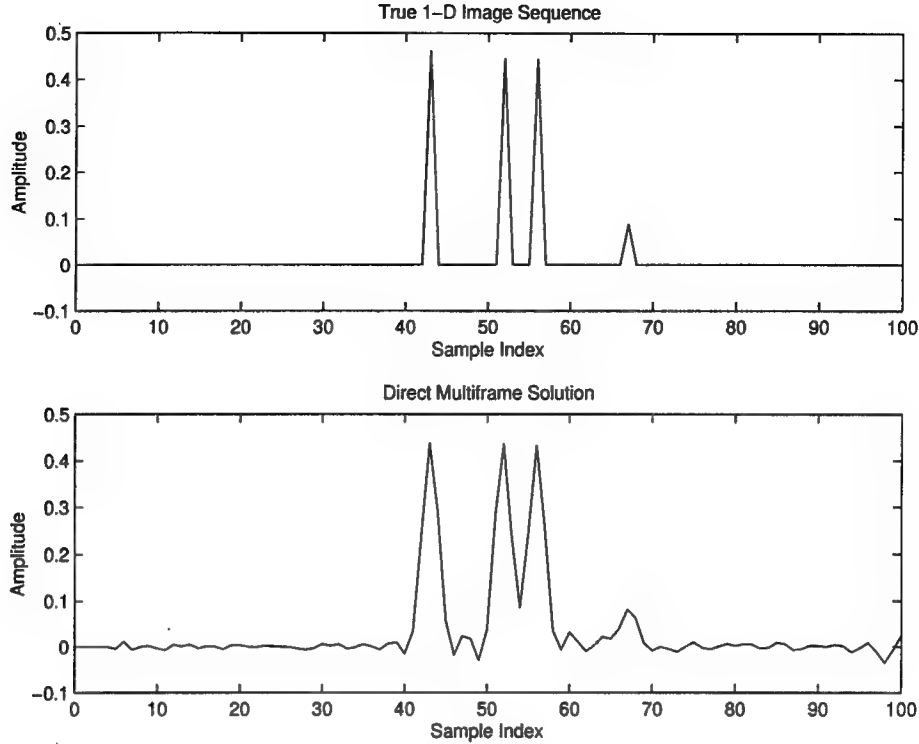


Figure 3: An example of the direct solution method for 1-D multiframe point source image restoration (cont.) a) [top] The true image sequence with four points (stars). b) [bottom] The estimated image sequence.

4 Cross Relationship Multiframe Blur Estimation

This section reports on studies involving direct estimation of the frame blur PSF's, \mathbf{H}_i , without jointly estimating the true image, \mathbf{f} . As with the direct signal (image) estimation method of Section 3, the direct blur methods are developed as adaptations of recently published algorithms found in the literature of digital communications, blind channel equalization [7, 8, 5, 6, 9]. By first obtaining high quality estimates of the blurs, it is then possible to use more conventional, well understood non-blind restoration algorithms to recover \mathbf{f} . This sequential approach (estimate \mathbf{H}_i first, then find \mathbf{f}) can yield higher accuracy signal estimates than a joint approach because the degrees of freedom in each estimation stage are reduced.

The cross relationship multiframe blur estimator exploits an algebraic relationship between frames to estimate PSF's, \mathbf{H}_i , independent of the object, \mathbf{f} . As shown in [2] and [9], the fact that the true image, \mathbf{f} , is common to every observation frame implies that in the noiseless case the following set of equations must hold

$$G_i = H_i * F \quad (9)$$

$$H_i * G_k - H_k * G_i = 0, \quad \forall i, k, i \neq k. \quad (10)$$

These have been called the "cross relationship" equations because each observed frame is "cross convolved" with another frame's unknown blur function. If the frame blurs have no common factors and \mathbf{f} contains sufficient modes (is persistently exciting) then the \mathbf{H}_i can be uniquely determined from equation (10). There are $(M^2 - M)/2$ equations in (10) when all combinations of i and k are evaluated, which in the absence of noise yields a fully determined system with a unique solution to within an arbitrary scale constant.

In [2, 9] a large matrix is formed which expresses all of the relationships of equations (10), and eigen-decomposition methods are then used to find the nullspace of this matrix. This nullspace contains the solution for the \mathbf{H}_i 's. This structure is not well suited to images because of the corresponding huge matrices for which eigen-decompositions must be computed. Additionally, these closed form solutions are not optimal in the presence of noise.

To address these problems, we have developed a computationally efficient iterative least squares algorithm that

is suitable for large image sizes. Equation (10) is approximated by

$$\hat{\mathcal{H}} = \arg \min_{\mathcal{H}} \sum_i \sum_{k \neq i}^{i-1} \|H_i * G_k - H_k * G_i\|_F^2, \quad (11)$$

where $\|\cdot\|_F$ indicates the Frobenius norm, and \mathcal{H} represents the set of all blur PSF frames, as in equation (5). For blurs with no common factors, equation (11) gives a unique solution for all H_i if an additional scale constraint is applied, and if F and the H_i are restricted to finite regions of support such that the entire convolution result is seen in G_i without truncation.

It is difficult to solve (11) jointly for all the H_i , so the proposed algorithm iterates by solving for each individual blur successively, while temporarily assuming the current blur estimates for the other frames are accurate. The iteration also cyclically applies region of support and scaling constraints. This approach of sequentially applying a series of constraints is essentially a projection onto convex sets (POCS) method.

The solution to equation (11) for the i^{th} blur frame, assuming all other H_k , $k \neq i$ are held fixed is given by

$$\frac{d}{dH_i} \sum_{k \neq i} \|H_i * G_k - H_k * G_i\|_F^2 = 2 \sum_{k \neq i} G_k * (G_k * H_i - G_i * H_k) = 0, \quad (12)$$

where \star indicates the 2-D deterministic correlation operation. In the frequency domain, (12) becomes the easily computed form

$$\tilde{H}_i = \frac{\tilde{G}_i \odot \sum_{k \neq i} \tilde{G}_k^* \odot \tilde{H}_k}{\sum_{k \neq i} \|\tilde{G}_k\|_F^2} \quad (13)$$

where \sim indicates the 2-D Fourier transform, * is the complex conjugate, and \odot is the element-by-element Shur matrix product.

The region of support and scaling constraints are more readily applied in the space domain, i.e. we require that

$$\mathcal{M} \odot H_i = 0, \text{ and} \quad (14)$$

$$\sum_{m,n} H_i(m,n) = 1.0, \quad (15)$$

where \mathcal{M} is a masking matrix with ones corresponding to the finite region of support for H_i , and zero elsewhere.

Combining these steps yields the following cross relationship blind blur estimation algorithm:

The Iterative Cross Relationship Blur Estimator

1. Compute the Fourier transform, $\tilde{G}_i = \mathcal{F}\{G_i\}$.
2. Iterate until done.
 - (a) Iterate for each frame, i .
 - i. Compute \tilde{H}_i using equation (eqn:crossfreq).
 - ii. Compute inverse Fourier Transform, $H_i = \mathcal{F}^{-1}\{\tilde{H}_i\}$
 - iii. Apply the region of support constraint, equation (14).
 - iv. Transform back to frequency domain, $\tilde{H}_i = \mathcal{F}\{H_i\}$.
 - (b) Apply the scaling constraint, equation (15).
3. End

Figure 4 illustrates results for this algorithm applied to five frames of synthetic AO star image data. Figure 4a is the original true image, and 4b shows the true blur frames, which are generated with an ellipsoidal Lorentzian function central peak, and a Markov random field mottled texture halo. The Lorentzian peaks have randomly generated major and minor axis length and orientation. This has been shown to be an excellent model for AO residual blur for the SOR telescopes [14, 15]. Figure 4c shows the observed data frames, which have noise added for a 50 dB PSNR level. Figure 4d shows the estimated blur frames. Note that the peak orientations are correctly estimated, however the random halo level is higher than in the true blurs.



Figure 4: Blind multiframe blur estimation example using cross relationship. a) Original point source image. b) True blur PSF's for five frames. c) Observed blurred image frames. d) Estimated blurs.

5 Cross Relationship Metric used in Iterative Bayesian Restoration

The following statistical approach to estimating the object and blur PSF's was initially developed during last year's (1997) AFOSR - SFRP tour. This approach has the advantage that available prior information regarding the object and the blur can be expressed in a natural way in the form of probability density functions (pdf's). These pdf's regularize the ill posed inverse problem, and can yield significant restoration performance improvement, particularly with lower SNR data sets. The generalized Gauss Markov random field (GGMRF) method is described here because during this year's tour (1998) some extensions to this work were developed, and the possibility of combining deterministic methods and these statistical methods was studied.

The cross relationship constraint of equation (11) represents additional available prior knowledge regarding the form of the solution. Any algorithm which does not exploit this information gives up potential performance improvement. We will express this information as a term in the blur PSF prior probability density function model used in Bayesian maximum *a posteriori* (MAP) image restoration. This term reduces the probability in the model for any solution which deviates significantly from the cross relationship constraint. The following sections review the development of the GGMRF blind restoration method, and shows how the method is adapted to incorporate the cross relation constraint.

5.1 The GGMRF Blind Bayesian Restoration Method

Assuming \mathbf{f} and \mathcal{H} are statistically independent (a reasonable assumption in practice), the blind MAP restoration problem may be stated as

$$\begin{aligned}\hat{\mathbf{f}}, \hat{\mathcal{H}} &= \arg \max_{\mathbf{f}, \mathcal{H}} p_{f,h|g}(\mathbf{f}, \mathcal{H}|\bar{\mathbf{g}}) \\ &= \arg \max_{\mathbf{f}, \mathcal{H}} p_{g|f,h}(\bar{\mathbf{g}}|\mathbf{f}, \mathcal{H}) p_f(\mathbf{f}) p_h(\mathcal{H})\end{aligned}\quad (16)$$

Models for the density functions in equation (16) are introduced here, and in the following two sections.

Assuming that the noise is white, Gaussian, and statistically independent from both the true image and the blur, the conditional probability of $\bar{\mathbf{g}}$ given \mathbf{f} and \mathcal{H} can be trivially shown to be

$$\begin{aligned}p_{g|f,h}(\bar{\mathbf{g}}|\mathbf{f}, \mathcal{H}) &= p_\eta(\bar{\mathbf{g}} - \mathcal{H}\mathbf{f}) \\ &= \frac{1}{\sqrt{2\pi}^{MN} \sigma_\eta} e^{-\frac{1}{2\sigma_\eta^2}(\bar{\mathbf{g}} - \mathcal{H}\mathbf{f})^t(\bar{\mathbf{g}} - \mathcal{H}\mathbf{f})}\end{aligned}\quad (17)$$

where N is the total number of pixels in the image field.

5.1.1 The GGMRF Image Prior Model

The image is modeled as a generalized Gaussian Markov random field (GGMRF), with joint pdf given by the Gibbs distribution [24]

$$p_f(\mathbf{f}) = \frac{1}{Z_f} \exp \left\{ -\beta \left[\sum_{\langle s,t \rangle \in \mathcal{C}_f} b_{s,t} |f_s - f_t|^p + \sum_{s \in S_f} a_s |f_s|^p \right] \right\} \quad (18)$$

where \mathcal{C}_f is the set of all cliques of pixels invoked by the specified neighborhood system, p is the GGMRF shape parameter, S_f is the set of all pixels in the image lattice, f_s denotes the s^{th} pixel in \mathbf{f} , and $b_{s,t}$ and a_s are potential weights. β is the scale parameter (controlling variance) and Z_f is the partition function, which is simply the scale constant required to insure the pdf integrates to 1.0.

The GGMRF model is a particularly flexible and useful for image restoration problems because it is capable of representing a wide variety of statistical image classes using just a few parameters. Equation (18) can be viewed as a definition for a random texture field. Parameters p , $b_{s,t}$ and a_s control the structure of this texture. p is known as the shape parameter, and controls the "edginess," or transition structure in the image. $b_{s,t}$ and a_s jointly control the correlation structure between neighboring pixels. With appropriate value selections, the model can be parameterized to generate fields that look like the detail in any desired visual texture, such as gravel, cork, grass, clouds, sand, waves, natural animal coat coloring, or geometric blocky structures.

Figure 5 shows an example of how shape parameter p affects this texture field. The images shown were generated as synthetic Markov random fields using Chen's algorithm to express the model of equation (18) [25]. A first order uniform neighborhood was used in each example, and only p was varied. We propose that Figure 5a, with $p = 0.5$, is a better statistical match for the blocky geometric structures of man-made objects, like satellites, than are the other three fields. Thus the GGMRF model with p values near 0.5 can serve as a powerful image prior to influence as restoration solution away from overly smoothed results typical of many other algorithms. This regularization is accomplished without detailed geometric descriptions of unknown object, but with a simple statistical preference for the desired texture-like pixel relationships.

For $.5 < p < 1.5$ (approximately) a MAP restoration using this GGMRF image prior preserves edge detail and regional structure found in most images of interest. The structural information contained in $p_f(\mathbf{f})$ is key to overcoming the convolutional ambiguity between \mathbf{f} and \mathbf{H} in blind restoration. Further, due to the Markov properties inherent in Gibbs distributions, the pdf of any single pixel depends only on the value of its neighbors as defined in the clique system, \mathcal{C} . This greatly simplifies the MAP optimization algorithm because updates need only consider local neighbor data rather than the entire image at each iteration.

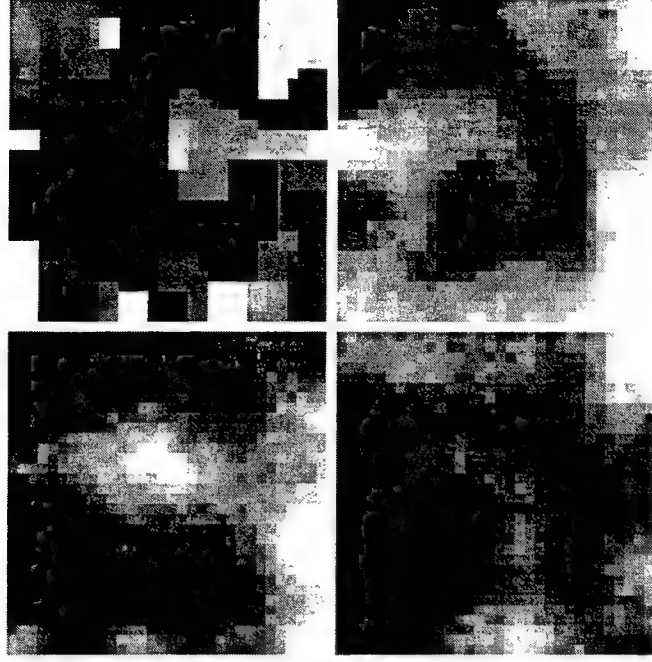


Figure 5: Synthetic GGMRF images. a) [top left] $p = 0.5$, b) [top right] $p = 1.0$, c) [bot. left] $p = 2.0$, d) [bot. right] $p = 3.0$.

5.1.2 The Composite GGMRF - Cross Relation Blur Model

The proposed blur pdf model need for blind Bayesian blur estimation is a composite of a GGMRF density and a Gibbs distribution based on the cross relationship of equation (11). The GGMRF is well suited to the task of modeling AO residual blur as well as the space object image because it can be adapted (by changing p) to favor the smoother, band-limited image types usually encountered as blur operators.

Our notions of what a blur field should look like are conveyed by specifying a few neighbor potential weights, the GGMRF shape parameter (q will be used to distinguish it from the object shaper parameter, p) and a mean value, μ_h . This in turn defines a texture class used to penalize the optimization objective function if a candidate blur differs significantly from the model.

The composite blur model is given by

$$p_h(\mathcal{H}) = \frac{1}{Z_h} \exp \left\{ -\alpha \left[\sum_{\langle s,t \rangle \in \mathcal{C}_h} c_{s,t} |(h_s - \mu_{h,s}) - (h_t - \mu_{h,t})|^q + \sum_{s \in S_h} d_s |h_s - \mu_{h,s}|^q \right] - \nu \sum_i \sum_{k \neq i}^{i-1} \|\mathbf{G}_i \mathbf{h}_k - \mathbf{G}_k \mathbf{h}_i\|^2 \right\} \quad (19)$$

where \mathcal{C}_h is the set of all cliques of pixels invoked by the specified neighborhood system, q is the GGMRF shape parameter, S_h is the set of all points in blur the lattice over all frames (set of every pixel in every frame), and $c_{s,t}$ and d_s are potential weights.

The first bracketed term of equation (19) represents the GGMRF portion of the blur pdf model. The terms on the second line is the cross relationship term. Under the assumption of white Gaussian noise, independent from frame to frame, the Gibbs distribution associated with the cross relationship term can properly be viewed as an independent Gaussian distribution on the residuals for the exact cross constraint of equation (10). It has been shown that for $1.75 < q < 3.0$ and \mathcal{C}_h corresponding to a first or second order uniformly weighted neighborhood, this model produces a random field with smoothed structures typical of many blur operators [26] [27].

5.2 Solution for GGMRF Blur and Object Models

In this section we present the MAP solution expression given the above models. A steepest descent algorithm to compute this solution is given in the following section. Substituting equations (19), (18) and (17) into (16), and performing a little algebra yields

$$\begin{aligned} \hat{\mathbf{f}}, \hat{\mathcal{H}} = & \arg \min_{\mathbf{f}, \mathcal{H}} \sum_{i=1}^M \|\mathbf{g}_i - \mathbf{H}_i \mathbf{f}\|^2 + \gamma \left[\sum_{\langle s, t \rangle \in \mathcal{C}_f} b_{s,t} |f_s - f_t|^p + \sum_{s \in S_f} a_s |f_s|^p \right] \\ & + \lambda \left[\sum_{\langle s, t \rangle \in \mathcal{C}_h} c_{s,t} |(h_s - \mu_{h,s}) - (h_t - \mu_{h,t})|^q + \sum_{s \in S_h} d_s |h_s - \mu_{h,s}|^q \right] \\ & + \vartheta \sum_i \sum_{k \neq i}^{i-1} \|\mathbf{G}_i \mathbf{h}_k - \mathbf{G}_k \mathbf{h}_i\|^2, \end{aligned} \quad (20)$$

where we have taken the logarithm of the r.h.s. (which does not change the maximization due to monotonicity) and have dropped additive constants. γ , λ and ϑ control the relative influence, as regularizing terms, that the image, GGMRF blur, and cross relationship pdf's have respectively on the solution.

The theoretically optimum values for γ , λ and ϑ are functions of p , q , and the SNR, though exact expressions for are difficult to derive. Therefor, these parameters are usually manually adjusted for most desirable results. The following proportionality relationships give some guidance on how to set γ and λ .

$$\gamma \propto 2\beta\sigma_\eta^2, \quad \lambda \propto 2\alpha\sigma_\eta^2 \quad (21)$$

Examples of restorations using the above formulation are presented in Section 5.4.

5.3 Steepest Descent Algorithm

For certain choices of the GGMRF parameters, equation (20) represents optimization of a convex functional, and can therefore be solved using the efficient gradient descent techniques (as compared with stochastic search methods needed for non-convex optimization). Sufficient conditions for the model parameters to insure convex optimization are: $q, p > 1$, $a_s, d_s \geq 0 \forall s$, and $b_{s,t}, c_{s,t} \geq 0 \forall s, t$. If $p = 2$, then a necessary and sufficient condition for the image prior term to be convex is that \mathbf{B} be a non-negative definite matrix with elements B_{st} related to equation (18) by $B_{st} = -b_{s,t}$, $a_s = \sum_{t \in \delta_s} B_{st}$, and $B_{st} = 0$ for $s \notin \delta_s$. δ_s is the set of neighboring pixels in all cliques containing f_s . Similar constraints hold for $c_{s,t}$ and d_s if $q = 2$. These conditions are not overly restrictive. We have found that with $p = 1.1$ and $q \geq 2$ the image model preserves edge features, and the blur model imposes sufficient smoothness to separate \mathbf{f} from \mathcal{H} .

In order to simplify gradient computation, the objective function of equation (20) can be expressed as

$$\begin{aligned} J(\mathbf{f}, \mathcal{H}) = & (\bar{\mathbf{g}} - \mathcal{H}\mathbf{f})^T (\bar{\mathbf{g}} - \mathcal{H}\mathbf{f}) + \gamma \text{Trace}\{\text{Diag}\{\sum_{t \in \delta_f} b_t |\mathbf{f} - \mathbf{P}_t \mathbf{f}|^p\}\} \\ & + \lambda \sum_i \text{Trace}\{\text{Diag}\{d |\mathbf{h}_i - \mu_h|^q + \sum_{t \in \delta_h} c_t |(\mathbf{h}_i - \mu_h) - \mathbf{P}_t (\mathbf{h}_i - \mu_h)|^q\}\} \\ & + \vartheta \sum_i \sum_{k \neq i}^{i-1} (\mathbf{G}_i \mathbf{h}_k - \mathbf{G}_k \mathbf{h}_i)^T (\mathbf{G}_i \mathbf{h}_k - \mathbf{G}_k \mathbf{h}_i) \end{aligned} \quad (22)$$

where $|\cdot|^p$ indicates element-wise exponentiation of the magnitude and $\text{Diag}\{\cdot\}$ forms a diagonal matrix from a vector. We have assumed $a_s = 0$ (a common choice [24]), and that $b_{s,t}$, $c_{s,t}$, and d_s are constant with respect to s . This assumption restricts the model to homogeneous (spatially stationary) random fields. \mathbf{P}_t is a permutation matrix which shifts \mathbf{f} to align f_t with f_s . δ_f and δ_h are sets of indices used to select shifts consistent with the neighborhood structure of \mathcal{C} . Since the field is stationary, these sets are independent of s . For example, in the case

of a first order neighborhood, δ_f would contain four indices to select up, down, left, and right neighbors. Neglecting edge effects, an example of \mathbf{P}_t where t = "nearest neighbor down" would be as follows:

$$\mathbf{P}_t = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ & \ddots & \ddots & & \\ 0 & \cdots & 0 & 1 & 0 \\ 0 & \cdots & & 0 & 1 \\ 0 & \cdots & & & 0 \end{bmatrix} \quad \text{when } t \text{ is the index for 1 row down.} \quad (23)$$

The derivative of equation (22) with respect to \mathbf{f} is readily computed. However, to compute the derivative with respect to the i^{th} frame blur, \mathbf{h}_i , the first term of (22) is replaced with the equivalent expression

$$\sum_{i=1}^M M(\bar{\mathbf{g}} - \mathbf{F}\mathbf{h}_i)^T (\bar{\mathbf{g}} - \mathbf{F}\mathbf{h}_i). \quad (24)$$

Computing gradients yields

$$\frac{d}{d\mathbf{f}} J(\mathbf{f}, \mathcal{H}) = -2 \sum_{i=1}^M \mathbf{H}_i^T (\mathbf{g}_i - \mathbf{H}_i \mathbf{f}) + \alpha p \sum_{t \in \delta_f} b_t \mathbf{Q}_t^T |\mathbf{Q}_t \mathbf{f}|^{p-1} \odot \text{sign}\{\mathbf{Q}_t \mathbf{f}\} \quad (25)$$

$$\begin{aligned} \frac{d}{d\mathbf{h}_i} J(\mathbf{f}, \mathcal{H}) = & -2\mathbf{F}^T (\mathbf{g}_i - \mathbf{F}\mathbf{h}_i) \\ & + \lambda q \left(\sum_{t \in \delta_h} c_t \mathbf{Q}_t^T |\mathbf{Q}_t (\mathbf{h}_i - \mu_h)|^{q-1} \odot \text{sign}\{\mathbf{Q}_t (\mathbf{h}_i - \mu_h)\} + d |\mathbf{h}_i - \mu_h|^{q-1} \odot \text{sign}\{\mathbf{h}_i - \mu_h\} \right) \\ & - 2\vartheta \sum_{k \neq i} (\mathbf{G}_k^T \mathbf{G}_i \mathbf{h}_k - \mathbf{G}_k^T \mathbf{G}_k \mathbf{h}_i), \end{aligned} \quad (26)$$

where $\mathbf{Q}_t = \mathbf{I} - \mathbf{P}_t$ and $\text{sign}\{\cdot\}$ is the signum function.

Having used matrix-vector notation to facilitate evaluating the gradients, we now note that the conventional matrix vector products in equations (25) and (26) involve huge matrices, and are extremely computationally inefficient. Fortunately, each product corresponds to either 2-D convolution or 2-D deterministic correlation in the image domain, which can be computed very efficiently in the frequency domain using fast Fourier transforms.

A simple and efficient iterative steepest descent algorithm based on the gradient expressions given above is

$$\mathbf{F}^{k+1} = \mathbf{F}^k + 2\alpha \sum_{i=1}^M \mathbf{H}_i^k \star (\mathbf{G}_i - \mathbf{H}_i^k \star \mathbf{F}^k) - \alpha \gamma p \sum_{t \in \delta_f} b_t \mathbf{Q}_t \star |\mathbf{Q}_t \star \mathbf{F}^k|^{p-1} \odot \text{sign}\{\mathbf{Q}_t \star \mathbf{F}^k\} \quad (27)$$

$$\begin{aligned} \mathbf{H}_i^{k+1} = & \mathbf{H}_i^k + 2\alpha \mathbf{F}^k \star (\mathbf{G}_i - \mathbf{F}^k \star \mathbf{H}_i^k) - \alpha \lambda q \sum_{t \in \delta_h} c_t \mathbf{Q}_t \star |\mathbf{Q}_t \star (\mathbf{H}_i^k - \mu_h)|^{q-1} \odot \text{sign}\{\mathbf{Q}_t \star (\mathbf{H}_i^k - \mu_h)\} \\ & - \alpha \lambda d |(\mathbf{H}_i^k - \mu_h)|^{q-1} \odot \text{sign}\{\mathbf{H}_i^k - \mu_h\} \\ & + \alpha 2\vartheta \sum_{k \neq i} \mathbf{G}_k \star (\mathbf{G}_i \star \mathbf{H}_k - \mathbf{G}_k \star \mathbf{H}_i), \end{aligned} \quad (28)$$

where all upper case variables represent 2-D images (rather than the column scanned images), superscripts, k , indicate the iteration number, ' \star ' represents 2-D deterministic correlation, ' \ast ' indicates 2-D convolution, and α is the iteration step size. The convolutions and correlations in equations (27) and (28) involving \mathbf{F} and \mathbf{H}_i can be computed efficiently in the frequency domain using a 2-D FFT based algorithm. Since \mathbf{Q}_t was formed as the sum of permutation and identity matrices, it is seen that convolutions and correlations involving \mathbf{Q}_t are computationally trivial, involving just a simple shift and add, i.e.

$$\begin{aligned} \mathbf{Q}_t \star \mathbf{F} &= \mathbf{F}[m, n] - \mathbf{F}[m - m_t, n - n_t], \quad [m - m_t, n - n_t] \in \{\text{the region of support for } \mathbf{F}\} \\ \mathbf{Q}_t \ast \mathbf{F} &= \mathbf{F}[m, n] - \mathbf{F}[m + m_t, n + n_t], \quad [m + m_t, n + n_t] \in \{\text{the region of support for } \mathbf{F}\} \end{aligned} \quad (29)$$

where m_t and n_t are the row and column distances to the neighbor indexed by t . Restoration results using this algorithm are presented in the following section.

5.4 Preliminary Experimental Results for GGMRF Restoration

The section presents some blind restoration results using real AO data, collected at the Starfire Optical Range (SOR) of the Air Force Research Laboratory. The SOR AO system includes a 1.5 meter telescope with a (approximately) 200 cell fully adaptive deformable mirror and a Shack-Hartmann wavefront detector of corresponding size in a closed loop control configuration. The data presented here is from a pass of a METEOR-1 satellite, collected in July or 1995. Partial AO correction was obtained using the object image itself at the wavefront detector, without the use of a guide star. Figure 6 is a sketch of the configuration of the METEOR-1 satellite class. In the results that follow, only the GGMRF portion of the blur model was used, i.e. $\vartheta = 0$. Experiments to evaluate performance with the cross relation term included are under way.

Figure 7 shows the data set used to compute the restoration. Each of the 10 frames shown is the average (without alignment correction) of 16 consecutive frames of the original data.

collected with a frame exposure time of 5 ms. This averaging increases SNR, albeit at the expense of the desired blur diversity. Figure 8 presents the steepest descent restoration achieved using GGMRF models for object and blur. Note that the solar panels are clearly resolved, and that shadowing at the junction of the lower panel and main body is evident. Figure 8b shows the result that would be obtained with no object regularization, i.e. without the GGMRF image prior term. This result is equivalent to the commonly used Maximum Likelihood solution. It is apparent that without the image prior model, the solution converges to an overly thinned, unrealistic region of support. Figure 8a is consistent with phase diversity processing of this data reported by Paxman et al [13], and with results using the algorithm of Jefferies and Christou [10] [11].

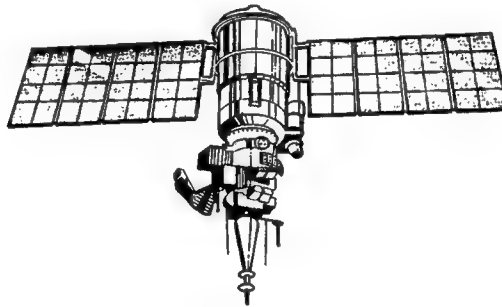


Figure 6: Diagram of the METEOR-1 satellite class. Drawing courtesy of USAF Phillips Laboratory [28].

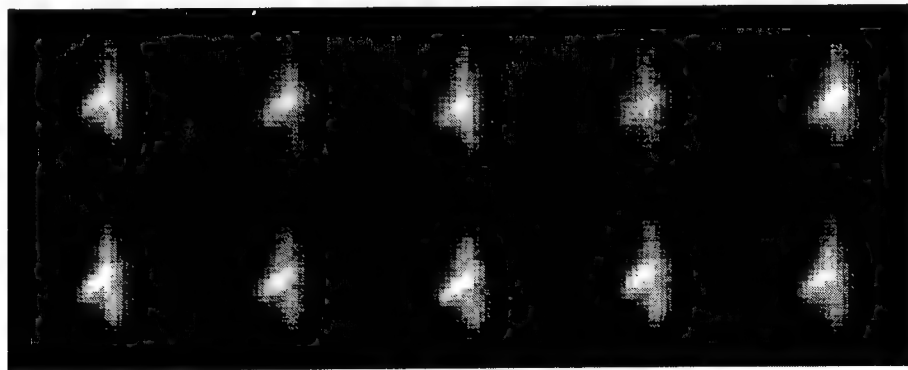


Figure 7: Observed data for catalog number 4419, METEOR-1 satellite. Each of the 10 frames consists of the average over 16 sequential original 5 ms exposure frames. Frames are 32x32 pixels. Note the shift in alignment and blur extent from frame to frame.



Figure 8: Restoration of the data in Figure 7. A logarithmic grayscale is used to emphasize low level detail. 6,000 iterations of the steepest descent algorithm were used, with $p = 1.1$, $q = 2.5$, first order GGMRF object model, and second order blur model. a) [left] Solution with $\gamma = .4$. b) [right] Solution with $\gamma = 0$, i.e. using no image prior model. This is equivalent to the Maximum likelihood solution, and is seen in this example to overly thin the solution.

6 Additional Accomplishments for the 1998 SFRP Tour

This section summarizes accomplishments resulting from the AFOSR Summer Faculty Research Fellowship Program support of my research collaboration with Starfire Optical Range at Kirtland Air Force Base. These activities are in addition to the research progress reported above, and include work completed from the end of the 1997 tour until the present.

Accomplishments:

1. Technical Papers Resulting From SOR Work:

- (a) Brian D. Jeffs and Julian Christou, "Blind Bayesian restoration of adaptive optics images using generalized Gaussian Markov random field models," *Proceedings of the SPIE, Conference 3353, Adaptive Optical System Technologies*, Mar. 23-26, Kona Hawaii.
- (b) Brian D. Jeffs, Julian Christou and Sheila Hong, "Blind MAP Restoration Of Adaptive Optics Telescope Images Using Estimated Shape Parameters For Generalized Gaussian MRF Models," *Proceedings of the 1998 IEEE DSP Workshop*, Bryce Canyon UT, Aug. 9-12, 1998.
- (c) Brian D. Jeffs, Sheila Hong, and Julian Christou, "A Generalized Gauss Markov Model for Space Objects in Blind Restoration of Adaptive Optics Telescope Images," 1998 IEEE International Conf. on Image Processing, Chicago IL, Oct. 4-7, 1998.
- (d) Brian D. Jeffs, Julian Christou, and Sheila Hong, "A Unified Model Based Approach for Blind Bayesian Restoration of Adaptive Optics Images," in preparation, to be submitted to *IEEE Transactions on Image Processing*, November 1998.

2. Master's thesis: Hong Hong, *Blind Bayesian Restoration for Adaptive Optics Telescope Images*, Brigham Young University, Aug. 1998.
3. Developed a data base of satellite images for space object model parameter estimation.
4. Student support: Have supported three graduate students as research assistants to pursue this topic for their theses: Sheila Hong, Brent Chipman and Miguel Apezteguia.
5. Developed an ongoing fruitful collaboration with lab point-of-contact, Julian Christou. Several joint projects are in progress.
6. Estimated GGMRF image prior model shape parameters for a variety of spacecraft. This has validated the approach, and helped improved restoration results.

7. Developed a method for GGMRF neighborhood influence parameter estimation. This also has helped improve restoration results.
8. Developed simulation software to synthesize realistic AO residual Blur Frames. This has simplified the process of evaluating algorithm changes.
9. Ongoing work (in addition to the report above) includes improving parameter settings for the GGMRF model for satellites and blur representation.

7 Conclusions

This report has detailed the progress made by the author during the 1998 AFOSR Summer Faculty Research Program for research in multiframe blind restoration of space objects images from adaptive optics systems. This effort has been very fruitful, and some very promising new technologies have been identified. It is expected that this research effort will lead to higher resolution restorations of AO images with less computational burden than other blind algorithms.

It is hoped that the deterministic and composite methods introduced above will enable better satellite structural detail resolution than is currently achievable in typical observing with the SOR 3.5 m AO system.

The early work on multiframe blind point source image restoration suggests that this problem will yield a breakthrough solution that will dramatically enhance system performance when computing parameters for binary stars or for dense star fields. The approach is quite different than any previously applied to the problem, and preliminary simulations look good.

This research program has also had a profound effect on three of my graduate students. Sheila Hong has recently graduated with a Master's degree, and a fine thesis reporting her work in AO image post processing. Also, Brent Chipman and Miguel Apezteguia have each recently started their thesis work based on this research program.

I expect that my research in this fruitful area and collaboration with Starfire Optical Range will continue for a number of years.

References

- [1] H. Liu and G. Xu, "A deterministic approach to blind symbol estimation", *IEEE Signal Processing Letters*, vol. 1, no. 12, pp. 205-207, Dec. 1994.
- [2] M. I. Gurelli and C. L. Nikias, "Evam: An eigenvector-based algorithm for multichannel blind deconvolution of input colored signals", *IEEE Trans. on Signal Processing*, vol. 43, no. 1, pp. 134-149, Jan. 1995.
- [3] L. Tong, G. Xu, B. Hassibi, and T. Kailath, "Blind channel identification based on second-order statistics: A frequency-domain approach", *IEEE Transactions on Information Theory*, vol. 41, no. 1, pp. 329-333, Jan. 1995.
- [4] E. Moulines, P. Duhamel, J. F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel fir filters", *IEEE Trans. on Signal Processing*, vol. 43, no. 2, pp. 516-525, Feb. 1995.
- [5] D. T. M. Slock, "Blind fractionally-spaced equalization, perfect-reconstruction filter banks and multichannel linear prediction", in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Adelaide, SA, Australia, Apr. 1994, vol. 4, pp. 585-588.
- [6] D. T. M. Slock and C. B. Papadakis, "Further results on blind identification and equalization of multiple fir channels", in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Detroit, MI, May 1995, vol. 3, pp. 1964-1967.
- [7] Y. Hua, "Fast maximum likelihood for blind identification of multiple fir channels", *IEEE Trans. on Signal Processing*, vol. 44, no. 3, pp. 661-672, Mar. 1996.

- [8] H. Liu, G. Xu, and L. Tong, "A deterministic approach to blind identification of multi-channel fir systems", in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Adelaide, SA, Australia, Apr. 1994, vol. 4, pp. 581-584.
- [9] G. Harikumar and Y. Bresler, "Efficient algorithms for the blind recovery of images blurred by multiple filters", in *Proceedings of ICIP-96, 1996 IEEE Int. Conf. on Image Processing*, Lausanne, Switzerland, Sept. 1996, vol. 3, pp. 97-100.
- [10] S.M. Jefferies and J.C. Christou, "Restoration of astrononmical images by iterative blind deconvolution", *Astrophysical Journal*, vol. 415, pp. 862-874, Oct. 1993.
- [11] J.C. Christou, D. Bonaccini, and Nancy Ageorges, "Deconvolution of adaptive optics near-infrared system images", *Proceedings of the SPIE*, vol. 3126: Conference on Adaptive Optics and Applications, July 1997.
- [12] JM. Conan, L. Laurent, T. Fusco, V. Michau, and G. Rousset, "Deconvolution of adaptive optics images using the object autocorrelation and positivity", *Proceedings of the SPIE*, vol. 3126: Conference on Adaptive Optics and Applications, July 1997.
- [13] R.B. Paxman, J.H. Seldin, B.J. Thelen, M.F. Reiley, and D.A. Carrarra, "Focused-beam scoring techniques, volume i", Tech. Rep. 278140-9-F(V.I), SOR, Air Force Phillips Laboratory, Kirtland AFB, 1997.
- [14] J.D. Drummond, "Sizes, shapes and rotational poles of ceres and vesta from adaptive optics images", *Bulletin of the American Astronomical Society*, vol. 27, pp. 16, 1996.
- [15] J.D. Drummond, R.Q. Fugate, J.C. Christou, and E.K. Hege, "Full adaptive optics imgaes of asteroids ceres and vesta: Rotational poles and triaxial ellipsoid dimensions", *to appear in ICARUS*, 1998.
- [16] R. L. Lagendijk and J. Biemond, *Iterative Identification and Restoration of Images*, Kluwer, Boston, 1991.
- [17] P.J. Sementilli, M.S. Nadar, and B.R. Hunt, "Poisson map super-resolution estimator with smooothness constraint", *Proceedings of the SPIE*, vol. 2032, July 1993.
- [18] Y. Yang, N.P. Galatsanos, and H. Stark, "Projection-based blind deconvolution", *Journal of the Optical Society of America A*, vol. 11, no. 9, Sept. 1994.
- [19] Z. Mou-yan and R. Unbehauen, "New algorithms of two-dimensioan blind deconvolution", *Optical Engineering*, vol. 34, no. 10, pp. 2945-2956, Oct. 1995.
- [20] A. K. Jain, *Fundamentals of Digital Image Processing*, Prentice Hall, Englewood Cliffs, 1989.
- [21] Alan V. Oppenheim and Ronald W. Schafer, *Discrete-Time Signal Processing*, Prentice Hall, Englewood Cliffs, New Jersey, 1989.
- [22] E. R. Cole, "The removal of unknown image blurs by homomorphic filtering, utec-csc-74-029", Tech. Rep., Department of Computer Science, University of Utah, 1973.
- [23] G. B. Giannakis and R. W. Heath Jr., "Blind identification of multichannel fir blurs and perfect image restortion", in *Proceedings of ICIP-96, 1996 IEEE Int. Conf. on Image Processing*, Lausanne, Switzerland, Sept. 1996, vol. 1, pp. 717-720.
- [24] C. Bouman and K. Sauer, "A generalized Gaussian image model for edge-preserving MAP estimation", *IEEE Trans. on Image Processing*, vol. 2, no. 3, pp. 296-310, July 1993.
- [25] C. C. Chen, *Markov random field models in image analysis*, PhD thesis, Michigan State University, East Lansing, 1988.
- [26] W. H. Pun and B. D. Jeffs, "Shape parameter estimation for generalized gaussian markov random field models used in map image restoration", in *Conference Record, 29th Asilomar Conf. Signals, Syst., Comp.*, Oct. 1995, pp. 1472-1476.

- [27] Wai Ho Pun, *Parametric Model-Adaptive Image Restoration*, PhD thesis, Brigham Young University, 1995.
- [28] N.L. Johnson and D.M. Rodvold, "Europe and asia in space, 1993-1994", Tech. Rep., USAF Phillips Laboratory/XPF, Kirtland AFB, NM, 1994.

SELF-SENSING TECHNIQUE FOR
ACTIVE ACOUSTIC ATTENUATION

Donald J. Leo
Assistant Professor
Mechanical Engineering Department

Virginia Polytechnic Institute and State University
Mechanical Engineering Department
Blacksburg, VA 24061-0261

Final Report for:
Summer Faculty Research Program
Air Force Research Lab - Phillips Site
Albuquerque, New Mexico

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Air Force Research Lab - Phillips Site

August 1998

SELF-SENSING TECHNIQUE FOR ACTIVE ACOUSTIC ATTENUATION

Donald J. Leo

Assistant Professor

Mechanical Engineering Department

Virginia Polytechnic Institute and State University

Douglas Limpert

Research Assistant

Mechanical, Industrial, and Manufacturing Engineering Department

The University of Toledo

ABSTRACT

A 'self-sensing' technique is developed for actively suppressing acoustic resonances of enclosed cavities. The technique enables the simultaneous measurement of mechanical velocity and acoustic pressure for a speaker driving an acoustic load. The technique is derived from an impedance model of the coupled electrical, mechanical, and acoustic systems. The velocity and pressure signals are obtained by passing the speaker voltage and current through linear filters that model the electrical and mechanical impedance of the speaker. A transfer function analysis of the technique demonstrates that the pressure and velocity measurements are sensitive to estimation of the physical parameters of the speaker, but that the technique is insensitive to variations in the acoustic load. Experimental results for a speaker driving a cylindrical cavity illustrate that accurate estimates of pressure are obtained for frequency ranges on the order of three octaves (20 - 150 Hz). Errors at higher frequencies are attributed to weaker than expected structural-acoustic coupling and inaccurate estimates of the electrical impedance. Active damping of acoustic resonances is demonstrated with second-order feedback compensators that utilize the self-sensing technique derived in the paper. Feedback control increases the damping in the first acoustic resonance from 4.4% critical to 14.9% critical and increases the damping in the second acoustic resonance from 2.4% critical to 3.2% critical. The acoustic response in the 40-60 Hz frequency range is increased due to small phase margins near the mechanical resonance of the speaker.

Table 1: List of Nomenclature

v_{in}	speaker input voltage	V
i	speaker current	A
e	voice coil back emf	V
R	electrical resistance	Ω
L	electrical inductance	H
z_a	acoustic mobility	$m^5/N\cdot s$
z_m	mechanical mobility	$m/N\cdot s$
Z_e	electrical impedance	Ω
Z_a	acoustic impedance	$N\cdot s/m^5$
Z_m	mechanical impedance	$N\cdot s/m$
f, f_1, f_2	mechanical force	N
p	acoustic pressure	N/m^2
u	speaker velocity	m/s
U	volume velocity	m^3/s
A	speaker area	m^2
Bl	voice coil force factor	Wb/m
M_{md}	speaker moving mass	kg
R_{ms}	speaker damping coefficient	$N\cdot s/m$
K_{ms}	speaker stiffness	N/m
ρ_o	density of air	kg/m^3
c	speed of sound	m/s
l	length of tube	m

Introduction

Active noise control has long been studied as a means of increasing the sound attenuation properties of materials and structures. The use of secondary sources for destructive interference of sound waves was experimentally demonstrated over 40 years ago [1, 2] but it wasn't until recently that advances in digital signal processing made active noise cancellation practical [3]. Continuing advances in digital signal processing technology have enabled the use of active noise control in several practical applications [4, 5].

A number of active control techniques utilize speakers and microphones as control elements. In the early work by Olson and May, an 'electronic sound absorber' was developed from a microphone placed in close proximity to the cone of an enclosed loudspeaker [1]. Reductions in the sound pressure were obtained by feeding back the pressure measurement through an analog circuit to the loudspeaker input. More recently, Nelson and his colleagues developed multi-channel control algorithms for active sound absorption and adaptive control implementations have been studied by Orduna-Bustamante and Nelson [6, 7, 8, 9]. In all of the works cited above, acoustic pressure and mechanical velocity were the variables utilized in the control algorithms. As pointed out by Elliot, *et al*, algorithms that utilize pressure and velocity as control variables result in power absorption. Optimal power absorption is achieved by matching the impedance of the electrical system to the impedance of the acoustic system with a feedback or feedforward compensator [8].

These works illustrate the utility of measuring acoustic pressure and velocity for active noise control. In this work, a 'self-sensing' technique is developed in which the acoustic pressure and velocity are measured simultaneously from measurements of speaker voltage and current. Unlike previous research, in which these quantities were measured using microphones and accelerometers, the technique derived in this paper enables the measurement of pressure and velocity without the need for additional sensors. In addition to eliminating the need for hardware in the control system, this technique enables the measurement of signals that are perfectly collocated with one another. Collocation has advantages in the development of dissipative feedback control laws.

The technique is motivated by previous research in the simultaneous sensing and actuation of flexible structures. The term 'self-sensing' was introduced by Dosch, Inman, and Garcia in their use of piezoelectric materials for the control of resonant modes in a flexible structure [10]. They demonstrated active feedback control of the first two modes of a flexible beam using a bridge circuit that simultaneously measured the input voltage and the mechanical strain in the material. A more detailed analysis of the electromechanical coupling was developed by Hagood, Anderson, and Goodliffe in their work on self-sensing actuation [11]. One of the detrimental aspects of the technique, 'feedthrough capacitance', was noted by both groups of researchers. The work by Dosch, *et al*, and Hagood, *et al*, was extended by Cole and Clark through the development of an adaptive filter algorithm that identified the feedthrough capacitance of the piezoceramic sensor-actuator [12].

More recently, the concept of self-sensing has been applied to applications in active noise control. Clark and Lane developed a technique for measuring the mechanical velocity of a speaker for the purpose of making it a constant volume velocity source over a specified frequency range [13]. They demonstrated that feedback control of the mechanical velocity eliminated instabilities caused by speaker dynamics and enabled the use of pressure measurements for suppressing acoustic resonances. The technique was applied to the active suppression of acoustic resonances within a reverberant enclosure [14].

The contribution of this work is the development of a technique for simultaneously measuring the mechanical velocity and acoustic pressure. Unlike the work by Lane and Clark, in which only the mechanical velocity was measured for the purpose of creating a constant volume velocity source, this work utilizes the pressure measurement directly in the control algorithm. Experimental results illustrate the use of the technique for suppressing the acoustic resonances of an enclosed cavity through the use of second-order feedback compensators.

The technique has applications in the development of robust control systems for active acoustic attenuation in launch vehicle payload fairings. This area is receiving increased attention due to the need for reducing the cost of placing delicate payloads into orbit. Active control has the potential for providing increased attenuation in frequency ranges in which passive attenuation is impractical. Vibroacoustic modeling of a particular launch vehicle payload fairing was performed by Leo, Anderson, and Glaese. Numerical results for a medium-sized shroud demonstrated the performance limitations of low-authority control algorithms utilizing structural sensors and actuators [15]. A more sophisticated approach to the problem of structural sensing was developed by Griffin, *et al*, using the concept of radiation modes for enclosed cavities excited by flexible structures [16].

The paper is organized in the following manner. First, impedance models are developed for the coupled electromechanical and mechanoacoustic problem. A transfer function analysis of the impedance model is then performed to investigate the sensitivity of the technique to measurement error. This investigation also yields insight into a method for tuning the algorithm such that maximum accuracy is achieved for both the velocity and pressure estimates. This estimation algorithm is then experimentally demonstrated and control

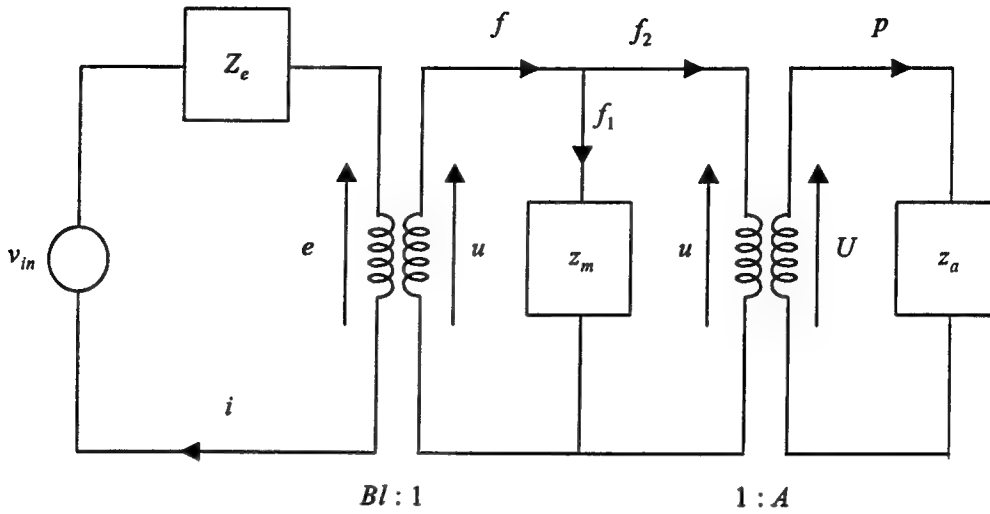


Figure 1: Electromechanical model of the speaker that incorporates the load of the acoustic element.

results are presented.

Impedance Modeling of the Speaker and Acoustic Load

The self-sensing technique is investigated with a system consisting of a closed cylindrical enclosure with a speaker mounted at one end. The control problem is represented by a three-mesh electromechanical impedance model, shown in Figure 1. The leftmost mesh corresponds to the electrical component, the middle mesh corresponds to the mechanical load of the speaker cone and voice coil, and the rightmost mesh corresponds to the acoustic load presented to the speaker [17]. The electrical and mechanical components of the system are a function of the speaker properties, while the acoustic load is a function of the type of acoustic element that is being driven by the speaker, e.g., an open room (modeled as a baffled acoustic element) or a tube with a specified length and diameter. Linear transformers model the electromechanical and mechanoacoustic coupling of the control system. The electromechanical transformer has a force factor of Bl , where B is the flux density of the permanent magnet and l is the length of wire that exists in the magnetic field. The mechanoacoustic transformer has a coupling coefficient that is a function of the speaker area A .

A set of equations that model the dynamics of the speaker and the acoustic load is derived from Figure 1. Summing current, velocity, and pressure in the meshes yields a set of three equations:

$$v_{in} + Z_e i - e = 0 \quad (1)$$

$$u + z_m f_1 = 0 \quad (2)$$

$$U + z_a p = 0 \quad (3)$$

Table 2: Definitions and parameters of the electromechanoacoustic model.

Parameter	Symbol	Value	Units
Force Factor	Bl	6.35	Wb/m
Moving Mass	M_{md}	0.04205	kg
Speaker Damping Coefficient	R_{ms}	1.21	N-s/m
Speaker Stiffness	K_{ms}	2998	N/m
density of air	ρ_o	1.18	kg/m ³
speed of sound	c	343	m/s
length of tube	l	1.22	m

Definitions and units for these parameters are listed in Tables 1 and 2. The equations for the electrical-to-mechanical transformer are:

$$\begin{aligned} Bli &= f \\ e &= Blu \end{aligned} \quad (4)$$

and the equations for the mechanical-to-acoustic transformer are

$$\begin{aligned} Au &= U \\ f_2 &= pA. \end{aligned} \quad (5)$$

Equations (1) through (5) are combined to obtain an expression for the electrical impedance, including the terms due to the mechanical and acoustic load. Solving equation (3) for the pressure and using the equivalences stated in equation (5) yields:

$$p = -U/z_a = -Au/z_a, \quad (6)$$

which can be rewritten to yield

$$f_2 = -A^2 \frac{u}{z_a}. \quad (7)$$

From the mesh diagram, we note that $f_1 = f - f_2 = f + A^2 \frac{u}{z_a}$. Substituting this expression into equation (2), along with the expression $f = Bli$, yields

$$u + z_m \left(Bli + A^2 \frac{u}{z_a} \right) = 0, \quad (8)$$

which can be rearranged to yield

$$\left(\frac{1}{z_m} + A^2 \frac{1}{z_a} \right) u + Bli = 0. \quad (9)$$

Solving this equation for the speaker velocity and noting that the mechanical and acoustic mobilities are simply the inverses of the mechanical and acoustic impedances, we can substitute the resulting expression into equation (1),

$$v_{in} + \left(Z_e + \frac{B^2 l^2}{Z_m + A^2 Z_a} \right) i = 0, \quad (10)$$

which relates the voltage and current across the speaker. As expected, the voltage-to-current relationship is a function of the electrical, mechanical, and acoustic impedances. The coupling between the three separate elements is a function of the relative magnitudes of the impedances as well as the coupling coefficients Bl and A .

Real-time Estimation of the Pressure and Velocity

The coupling inherent in the impedance model enables us to measure the speaker voltage and current and estimate the mechanical velocity and acoustic pressure. Assuming that we can measure both the voltage, v_{in} , and current, i , we can estimate the mechanical velocity from the expression:

$$\hat{u} = \frac{e}{Bl} = \frac{1}{Bl} (v_{in} + \hat{Z}_e i), \quad (11)$$

where \hat{Z}_e is an estimate of the electrical circuit impedance. Combining equations (10) and (11) yields

$$\frac{\hat{u}}{i} = \frac{\hat{Z}_e - Z_e}{Bl} - \frac{Bl}{Z_m + A^2 Z_a}. \quad (12)$$

If the estimated electrical impedance is exactly equal to the actual electrical impedance, then the estimated velocity is equivalent to the speaker velocity.

The expression for the estimated pressure can be derived in a similar manner. Substituting $f_1 = Bli - pA$ into equation (2) and multiplying through by the estimated mechanical impedance, \hat{Z}_m , yields the following expression

$$\hat{Z}_m u + Bli = pA. \quad (13)$$

Again assuming that we can measure the current across the speaker, and assuming that we use the estimated velocity from equation (12), we can estimate the pressure from the expression

$$\hat{p} = \frac{1}{A} (\hat{Z}_m \hat{u} + Bli), \quad (14)$$

where \hat{Z}_m is the estimated mechanical impedance.

Pole-Zero Models for the Speaker and Acoustic Load

Equations (12) and (14) represent a method for estimating the velocity and pressure without directly measuring either one of these quantities. The primary advantages of indirectly estimating the pressure and velocity are the elimination of any additional sensors and the reduction of measurement errors due to transducer dynamics.

The equivalent circuit model for the speaker can be found in several different handbooks and textbooks on acoustics and sound reproduction (see, for example, Beranek [17]). A standard model for the blocked electrical impedance of a speaker is

$$Z_e = \frac{v_{in}}{i} = Ls + R, \quad (15)$$

where L is the speaker inductance, R represents the speaker resistance, and s is the Laplace variable. We will see shortly that this model for the electrical impedance is not accurate enough for our current investigation, and a more accurate model will be proposed in the next section.

Continuing with our analysis, if we assume that the speaker cone is modeled as a simple mass-spring-damper, we can express the mechanical impedance in the Laplace Domain as

$$Z_m = \frac{M_{md}s^2 + R_{ms}s + K_{ms}}{s}, \quad (16)$$

where M_{md} , R_{ms} , and K_{ms} are the mechanical mass, damping, and stiffness, respectively.

The tube is modeled as closed at both ends, therefore the acoustic impedance at the end driven by the speaker is

$$Z_a = -j(\rho_o c/A) \cot(kl), \quad (17)$$

where ρ_o is the density of air and c is the speed of sound. The wavenumber k is equivalent to $\frac{\omega}{c}$, where ω is the angular frequency, and l is the effective length of the tube. This model for the acoustic impedance assumes that there is no energy dissipation in the tube.

A pole-zero model of the acoustic impedance can be developed from equation (17) by computing the frequencies at which $Z_a = 0$ and $Z_a \rightarrow \infty$. These are:

$$\begin{aligned} Z_a = 0 & \quad \omega_{zi} = (2i-1) \frac{\pi c}{2l} \quad i = 1, 2, \dots \\ Z_a \rightarrow \infty & \quad \omega_{pi} = i \frac{\pi c}{l} \quad i = 1, 2, \dots \end{aligned} \quad (18)$$

Using the previous expressions, we can write an N -term pole-zero model of the acoustic impedance as

$$Z_a = \frac{1}{s} \frac{\rho_o c^2}{Al} \frac{\prod_{i=1}^N (s^2/\omega_{zi}^2 + 1)}{\prod_{i=1}^N (s^2/\omega_{pi}^2 + 1)}. \quad (19)$$

The pole-zero models of the acoustic impedance, mechanical impedance, and electrical impedance are combined to form a single model of the speaker and tube. This model will be used to analyze the dynamics of the system and determine the robustness of the real-time estimation technique.

Transfer Function Analysis

Assuming perfect estimation of the mechanical and electrical impedance, equations (12) and (14) will enable the estimation of the speaker velocity and applied pressure. Perfect estimation is never possible, and therefore a transfer function analysis is performed to determine the sensitivity of the technique to imperfect estimation of the mechanical and electrical impedance.

If we denote

$$N_e(s) = Ls + R \quad \hat{N}_e(s) = \hat{L}s + \hat{R}, \quad (20)$$

and,

$$N_m(s) = M_{md}s^2 + R_{ms}s + K_{ms} \quad \hat{N}_m(s) = \hat{M}_{md}s^2 + \hat{R}_{ms}s + \hat{K}_{ms}, \quad (21)$$

and assume that,

$$Z_a = \frac{1}{s} \frac{\rho_o c^2}{Al} \frac{\prod_{i=1}^N (s^2/\omega_{zi}^2 + 1)}{\prod_{i=1}^N (s^2/\omega_{pi}^2 + 1)} = \frac{1}{s} \frac{\rho_o c^2}{Al} \frac{N_a(s)}{D_a(s)}, \quad (22)$$

we can write

$$Z_e = \frac{N_e(s)}{1} \quad \hat{Z}_e = \frac{\hat{N}_e(s)}{1} \quad Z_m = \frac{N_m(s)}{s} \quad \hat{Z}_m = \frac{\hat{N}_m(s)}{s} \quad Z_a = \frac{1}{s} \frac{\rho_o c^2}{Al} \frac{N_a(s)}{D_a(s)}. \quad (23)$$

The expression for the velocity estimate, equation (12), can be rewritten

$$\frac{\hat{u}}{i} = \frac{\hat{N}_e(s) - N_e(s)}{Bl} - \frac{sBlD_a(s)}{P(s)} \quad (24)$$

where $P(s) = N_m(s)D_a(s) + k_{ma}^2 N_a(s)$ and $k_{ma}^2 = \rho_o c^2 A/l$ is the mechanical-to-acoustic coupling of the system. This expression is combined to yield

$$\frac{\hat{u}}{i} = \frac{[\hat{N}_e(s) - N_e(s)] P(s) - s(Bl)^2 D_a(s)}{BlP(s)} = \frac{\hat{N}_u}{BlP(s)}. \quad (25)$$

We can obtain a similar expression for the pressure estimate by dividing equation (14) by i and substituting the previous equation for \hat{u}/i :

$$\frac{\hat{p}}{i} = \frac{\hat{N}_m(s)}{s} \frac{\hat{N}_u(s)}{BlP(s)} + Bl. \quad (26)$$

which is combined to yield

$$\frac{\hat{p}}{i} = \frac{\hat{N}_m(s)\hat{N}_u(s) + s(Bl)^2 P(s)}{sBlP(s)}. \quad (27)$$

Equations (25) and (27) illustrate that errors in the estimate of the electrical and mechanical impedance affect only the *zeros* of the transfer functions between the current and the estimated quantity. The system poles defined by $P(s)$ are unaffected by errors in the impedance estimates.

The velocity and pressure estimates are related to the input voltage through equation (10). Rewriting this expression in terms of the numerator and denominators of the impedance functions yields

$$\frac{v_{in}}{i} = -\frac{N_e(s)P(s) + s(Bl)^2 D_a(s)}{P(s)}, \quad (28)$$

which can be inverted and multiplied by equations (25) and (26) to obtain

$$\begin{aligned} \frac{\hat{u}}{v_{in}} &= \frac{[\hat{N}_e(s) - N_e(s)]P(s) - s(Bl)^2 D_a(s)}{Bl[N_e(s)P(s) + s(Bl)^2 D_a(s)]} \\ \frac{\hat{p}}{v_{in}} &= \frac{\hat{N}_m(s)\hat{N}_u(s) + s(Bl)^2 P(s)}{sBl[N_e(s)P(s) + s(Bl)^2 D_a(s)]} \end{aligned} \quad (29)$$

Again note that the poles of the system, which are the roots of equation $N_e(s)P(s) + s(Bl)^2 D_a(s)$, are unaffected by errors in the estimated impedance; only the zeros of the transfer functions are a function of the estimated electrical and mechanical impedance. The negative sign shown in equation (28) has been neglected in the previous expression due to the fact that it only affects the phase of the estimated quantity. The transfer function model of the system will be used to study the effects of mechanical-to-acoustic coupling on the dynamics and the effects of estimation errors on the quantities \hat{u} and \hat{p} .

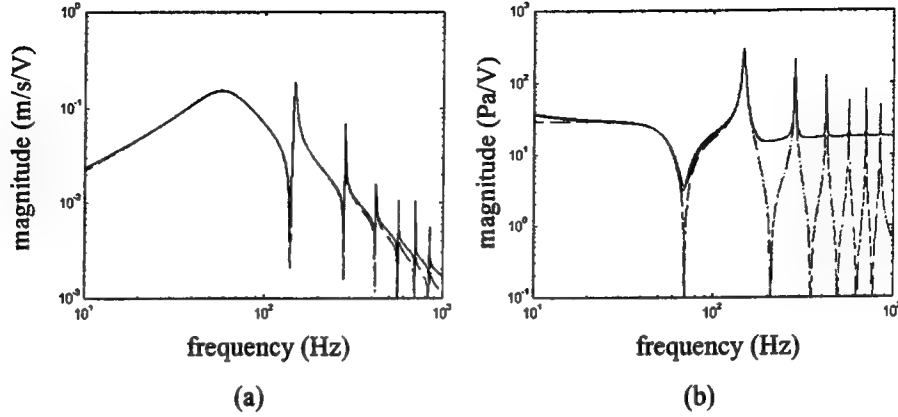


Figure 2: Effect of imperfect estimation of the electrical resistance: (a) mechanical velocity, and, (b) acoustic pressure.

Perfect Estimation of the Mechanical and Electrical Impedance

In the case where the electrical and mechanical impedances are known exactly, $\hat{N}_e(s) = N_e(s)$ and $\hat{N}_m(s) = N_m(s)$, the expressions in (29) reduce to

$$\begin{aligned} \frac{\hat{u}}{v_{in}} &= \frac{-s(Bl)D_a(s)}{N_e(s)P(s) + s(Bl)^2 D_a(s)} \\ \frac{\hat{p}}{v_{in}} &= Bl \frac{P(s) - N_m(s)D_a(s)}{N_e(s)P(s) + s(Bl)^2 D_a(s)} \end{aligned} \quad (30)$$

For a moment let us assume that the mechanical-to-acoustic coupling coefficient, k_{ma} , is equal to zero. In this case, $P(s) = N_m(s)D_a(s) + k_{ma}^2 N_a(s)$ simply reduces to $P(s) = N_m(s)D_a(s)$, which is substituted into the previous expression to yield

$$\begin{aligned} \frac{\hat{u}}{v_{in}} &= \frac{-s(Bl)}{N_e(s)N_m(s) + s(Bl)^2} = \frac{-s(Bl)}{N_e(s)N_m(s) + s(Bl)^2} \\ \frac{\hat{p}}{v_{in}} &= Bl \frac{N_m(s) - N_m(s)}{N_e(s)N_m(s) + s(Bl)^2} = 0. \end{aligned} \quad (31)$$

The fact that $\hat{p}/v_{in} = 0$ is consistent with the assumption that the mechanical-to-acoustic coupling coefficient is equal to zero. Also, the poles of the transfer function between the speaker velocity and speaker voltage are simply the roots of the equation $N_e(s)N_m(s) + s(Bl)^2 = 0$, which, as expected, demonstrates the dynamics of the acoustic impedance have no effect on the speaker dynamics when $k_{ma} = 0$.

Imperfect Estimation of the Mechanical and Electrical Impedance

Imperfect estimation of the mechanical and electrical impedance will affect the zeros of the transfer functions stated in equation (29). A numerical analysis was performed using the parameters shown in Table 2 to determine the sensitivity of the estimated frequency responses to imperfect impedance measurements. In the first case, it was assumed that the inductance and mechanical impedance estimates were exact and that the resistance was incorrect. The resistance was varied from 0.95 to 1.05 of its actual value (as stated in

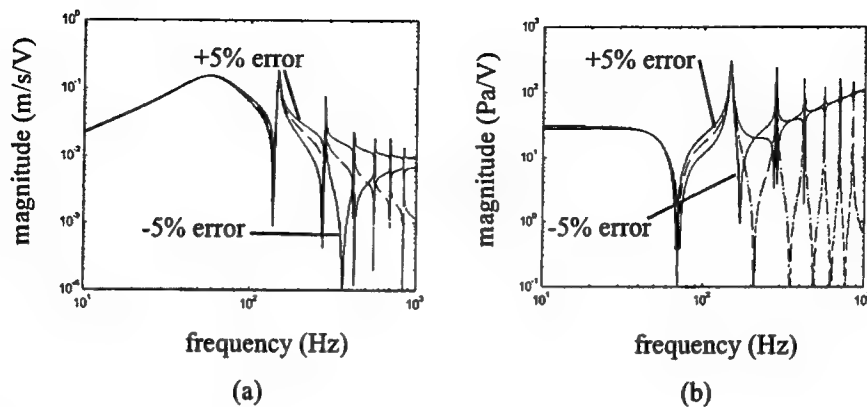


Figure 3: Effect of imperfect estimation of the electrical inductance: (a) mechanical velocity, and, (b) acoustic pressure.

Table 2) and the transfer functions stated in equations 29 were computed at each value of \hat{R} . Figure 2 illustrates the effect of varying the resistance on the estimate of the speaker velocity and applied pressure.

The numerical analysis demonstrates that imperfect estimation of the electrical resistance has a greater effect on the estimated pressure than the estimated velocity. The estimated velocity shows a slight increase in magnitude at higher frequencies (as compared to the actual velocity) for a $\pm 5\%$ error in the estimated resistance. In contrast, the pressure estimate shows significant error for the same values of \hat{R} . As Figure 2b illustrates, a $\pm 5\%$ error in the estimated resistance causes significant error in the zeros of the transfer function. The depth of the zeros, and hence the damping, shows considerable error as compared to the actual transfer function between the speaker voltage and the acoustic pressure.

A similar analysis was performed for the estimated inductance. The value of the electrical resistance and mechanical impedance was assumed to be equal to the actual values and the inductance varied between $0.95L$ and $1.05L$. The results are shown in Figure 3. Imperfect estimation of the electrical inductance has a considerable effect on both the estimated velocity and estimated pressure. Figure 3a illustrates that a -5% error in the electrical inductance creates a lightly damped zero in the transfer function between the input voltage and estimated velocity. A $+5\%$ error moves the zeros and poles closer together at higher frequencies, thus making acoustic resonances less prominent in the magnitude response.

Errors in the estimated inductance also have a considerable impact on the estimated pressure. A $\pm 5\%$ error moves the zeros and poles closer together, particularly at high frequencies, but does not significantly change the depth of the zeros. A -5% error reduces the frequency of each of the zeros and moves them closer to acoustic resonances. Likewise, a $+5\%$ error in the estimated inductance increases the frequency of each of the zeros and moves them closer to the poles of the transfer function. In either case, errors in the estimated inductance decrease the pole-zero spacing as compared to the actual transfer function.

The effect of imperfect stiffness measurements were also investigated numerically. In this case, it was assumed that the electrical impedance was measured perfectly, i.e., $\hat{R} = R$ and $\hat{L} = L$, but the stiffness varied $\pm 5\%$ from the actual value. The effect of imperfect stiffness measurements on the pressure estimate

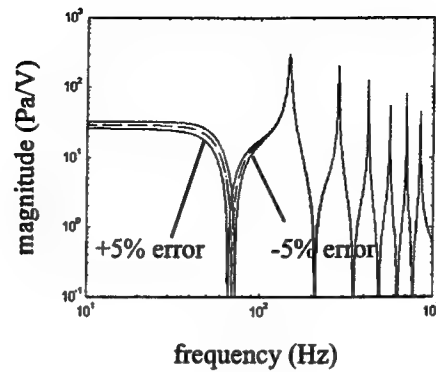


Figure 4: Effect of imperfect measurement of the mechanical impedance on the estimated pressure.

are shown in Figure 4. As the figure illustrates, imperfect estimation of the mechanical impedance only effects the pressure measurement in the region near the lowest frequency zero. As expected from the transfer function analysis, the poles are unaffected. Also unaffected is the estimate of the speaker velocity. As shown in equation (29), the estimate of the speaker velocity is not a function of the mechanical impedance, therefore it would not affect the transfer function between \hat{u} and v_{in} .

Summary of the Sensitivity Analysis

The sensitivity analysis demonstrates that inaccuracies in the mechanical and electrical impedance lead to different types of errors in the velocity and pressure estimates. Errors in the electrical resistance do not affect the velocity estimate significantly, but cause significant errors in the damping of the zeros of \hat{p}/v_{in} . Errors in the estimate of the electrical inductance cause zeros to appear in the velocity estimate and significantly change the frequency of the zeros of \hat{p}/v_{in} . The damping and frequency of the zeros introduced in the transfer function depend on whether the estimated inductance is greater or less than the actual inductance. In the final case study, errors in the estimate of the mechanical impedance caused no errors in the velocity estimate (as predicted by equation (29)), but did cause errors in the frequency region near the first zero of the transfer function \hat{p}/v_{in} . The results of this sensitivity analysis will be used in the tuning of the velocity and pressure estimates.

Experimental Results

The estimation technique is experimentally tested on a cylindrical enclosure with a speaker mounted at one end. One end of the tube is essentially a rigid end condition made from plywood while a plywood enclosure containing a 6 inch diameter speaker is mounted at the opposite end. An accelerometer is affixed with epoxy to the speaker cone and a microphone is placed on a bracket approximately 2 inches away from the cone surface. A second microphone is placed on the outer diameter of the rigid end of the tube. The wires for these transducers are connected through the plywood enclosures to eliminate holes in the enclosed cavity.

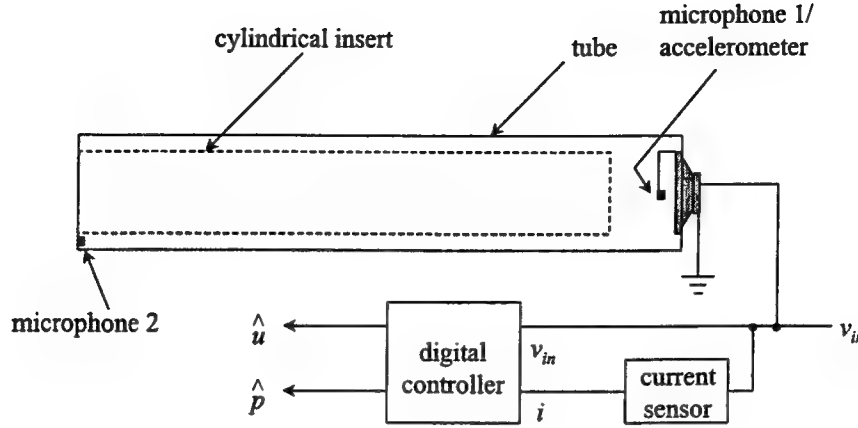


Figure 5: Experimental setup.

Signal conditioning is required for the accelerometer, microphones, and current sensor. The accelerometer is a standard ICP device from PCB Piezotronics. The pressure sensors are electret condenser microphones that require a simple resistor-capacitor network for signal conditioning. Current through the speaker is measured with a differential amplifier that measures the voltage drop across a 0.61Ω resistor. Data acquisition and control systems are used for dynamic analysis and real-time implementation of the estimation technique. An eight-channel spectral analyzer measures input-output transfer functions and a five-input, six-output digital signal processing system is used for real-time estimation and control.

As discussed in the introduction, one of the advantages of the technique developed in this paper is an insensitivity to the acoustic load. To test this concept experimentally, the rigid end of the tube is designed such that a second cylindrical section can be placed inside of the tube. The interior cylinder is made from PVC tubing and includes an end cap so that it is a closed volume. Its geometry is designed such that approximately 70% of the interior volume of the enclosure is occupied by the PVC tubing. Placing the PVC tubing inside of the enclosure has a significant effect on the acoustic resonances of the enclosed volume and enables us to test the sensitivity of the estimation technique to changes in the acoustic load.

The first set of tests is performed on the speaker mounted outside of the enclosure and radiating into a large open room. This test is performed to determine the electromechanical impedance of the speaker for a small acoustic load. Assuming that the acoustic load is small enough such that $|A^2 Z_a| \ll |Z_m|$, the combined electrical-mechanical impedance is approximately

$$\frac{v_{in}}{i} \approx Z_e + \frac{(Bl)^2}{Z_m}. \quad (32)$$

The results of the impedance test are shown in Figure 6. As expected, the impedance is purely real at low frequencies but exhibits a peak near the mechanical resonance of the speaker.

The only anomaly in the test results is slope of the impedance magnitude at frequencies above approximately 150 Hz. Under the assumption that the electrical impedance is modeled as the transfer function $Z_e = Ls + R$, it's clear that the slope of the impedance magnitude on a log-log plot should simply be 1.

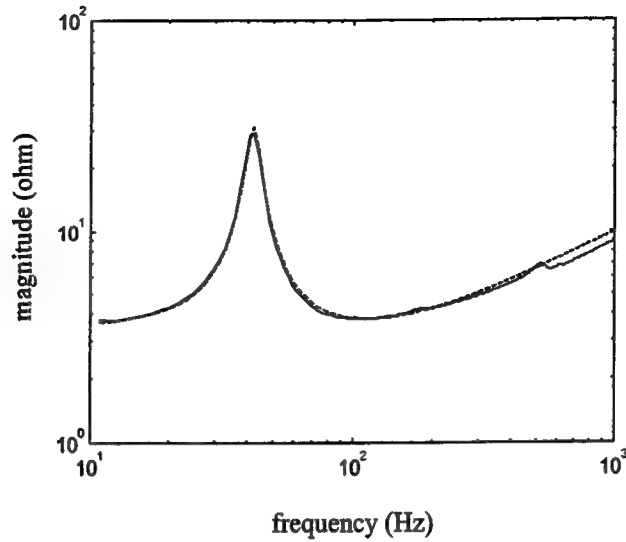


Figure 6: Measured speaker impedance radiating into an open room (solid) and the magnitude of the pole-zero impedance model (dashed).

Curvefitting the measured data reveals that the log-log slope of the impedance magnitude is 0.602.

The noninteger slope exhibited by the measured data requires a change in the model of the electrical impedance. As the numerical sensitivity analysis demonstrated, precise matching of the electrical and mechanical impedance is required to achieve accurate pressure and velocity estimates. Thus, a simple inductive model of the electrical impedance is not accurate enough for our purposes. An interlaced pole-zero model of the form

$$\hat{Z}_e = \hat{R} \prod_{m=-N}^N \frac{\frac{\hat{L}}{\hat{R}} \Delta^{N-m} s + 1}{\frac{\hat{L}}{\hat{R}} \Delta^{N-m+\gamma/2} s + 1} \quad (33)$$

is utilized instead. From the standpoint of modeling the electrical impedance, a pole-zero model of the form has several beneficial characteristics. The dc gain of the system is simply \hat{R} due to the form of the numerator and denominator. For $\Delta < 1$, the transfer function has a positive slope of $\gamma/2$ at high frequencies and the first transfer function zero occurs at $\omega = \hat{L}/\hat{R}$. The parameter Δ is chosen to make the transfer function estimate accurate over approximately a decade, i.e., 150 to 1500 Hz. With these parameters, the estimated electrical impedance is expressed by

$$\hat{Z}_e = \hat{R} \prod_{m=-5}^5 \frac{\frac{\hat{L}}{\hat{R}} (0.691)^{5-m} s + 1}{\frac{\hat{L}}{\hat{R}} (0.691)^{5-m+1.204} s + 1} \quad (34)$$

A comparison of the estimated impedances and the measured values are also shown in Figure 6. The interlaced pole-zero model is able to accurately represent the high frequency behavior of the electrical impedance.

Pressure Estimation

The pressure estimation technique is tested by placing the speaker in the enclosure and comparing the estimated pressure with a direct measurement of the quantity. The equations for estimated pressure and estimated velocity are implemented in real time on a digital signal processor sampling at 28.5 kHz. The

sampling rate is deemed high enough compared to the bandwidth of interest (≈ 500 Hz) that the system is modeled in continuous time.

The block diagram of the digital implementation is shown in Figure 7. The current and voltage across the speaker is measured with an analog circuit and these two quantities are input to the digital signal processor. In accordance with equation (12), the velocity estimate is formed from the following expression:

$$\hat{u} = \frac{1}{Bl} (g_u v_{in} + \hat{Z}_e i) \quad (35)$$

The variable g_u is a gain that accounts for the error in the calibration of the current sensor. It is varied to obtain more accurate velocity estimates.

The pressure estimate was implemented by the following expression

$$\hat{p} = \hat{Z}_m \hat{u} + g_p (Bl) i, \quad (36)$$

which, again, contains a variable gain g_p to obtain a more accurate estimate of the pressure. Furthermore, to eliminate overflow in the processor, this pressure estimate does not contain the multiplicative factor $1/A^2$.

The pressure and velocity estimates are 'tuned' by changing the values of \hat{R} , g_u , \hat{L} , and g_p . At each parameter value, the frequency response between the estimated quantities and the input voltage (i.e., \hat{u}/v_{in} and \hat{p}/v_{in}) is compared to the frequency responses obtained with the accelerometer and microphone 1. The velocity measurement is obtained by dividing the acceleration response by $j\omega$, where ω is the measurement frequency in rad/sec.

Experimental results indicate that the process can be broken down into two steps. First, the parameters \hat{R} and g_u are chosen such that an accurate velocity estimate is obtained. The second step in the process involves varying \hat{L} and g_p to obtain an accurate estimate of the pressure. It is found experimentally that accurate estimates of the velocity lead to accurate pressure estimates in the low-frequency region of the response. Once \hat{R} and g_u are chosen properly, the region between approximately 20 and 80 Hz is accurate in the pressure response, but the frequency region above 80 Hz is very sensitive to the choice of \hat{L} and g_p . Figure 8 illustrates the effect of varying \hat{L} on the estimated pressure and velocity. As predicted by the numerical sensitivity analysis, inaccurate values of \hat{L} lead to errors in the frequency region near the acoustic resonances.

The frequency response between the input voltage and the pressure estimates are shown in Figure 9 for the best choice of the tuning parameters. The results demonstrate that an accurate estimate of the pressure, both in magnitude and phase, is obtained over the frequency range 20 to 150 Hz. Above this frequency the pressure estimate deviates substantially from the actual pressure response. The deviation causes all of the acoustic resonances above 150 Hz to be unobservable in the frequency response of the pressure estimate. The errors in the frequency response above 150 Hz are attributed to two factors. First, the mechanical-to-acoustic coupling between the speaker and the internal resonances is not as great as the model predicts. The weaker-than-expected coupling is attributed to the fact that the speaker does not completely fill the interior cavity of the enclosure. Thus, one of the primary assumptions of the model is violated. The geometry of the speaker makes the sound radiation somewhere between that of a baffled piston and one which completely fills the end of the enclosure, therefore it would be expected that the mechanical-to-acoustic coupling would be weaker than predicted by the analytical model. The weaker-than-expected coupling makes the pressure estimate more sensitive to errors in the estimated parameters. The deviation between the pressure estimate and actual response at frequencies above 150 Hz is attributed to small errors in the estimate of the high frequency gain and phase of the speaker impedance.

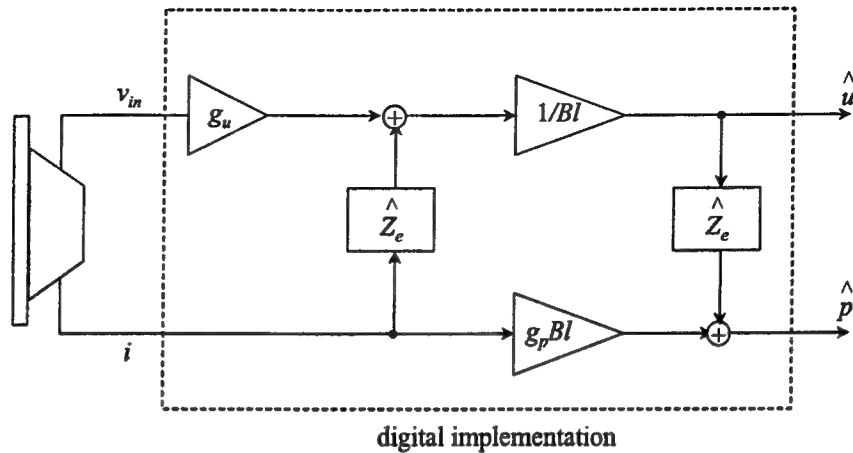


Figure 7: Block diagram of the digital implementation of the pressure and velocity estimation technique.

The sensitivity of the pressure estimation technique to changes in the acoustic load is tested by placing the PVC cylinder inside the tube and measuring the pressure estimate. The cylinder on the interior of the tube is made from plastic and is capped to eliminate any Helmholtz resonator effects. The geometry of the cylinder is such that it takes up approximately 70% of the interior volume of the tube and leaves a small gap (≈ 1.75 cm) between the outer diameter of the cylinder and the outer diameter of the tube. The 'fill factor' is defined as the ratio of volume occupied to the total volume of the empty tube.

The size of the interior cylinder significantly affects the acoustic resonances of the interior cavity. For the empty tube, the acoustic resonances are approximately integer multiples of 140 Hz, as predicted by the numerical analysis. With the interior cylinder placed in the tube, the first acoustic resonance decreases to 100 Hz and a second mode appears at 132 Hz.

The results of the pressure estimation for the case of 70% fill factor are also shown in Figure 9. The results demonstrate that the pressure estimate is accurate over the same frequency range as the case of zero fill factor (i.e., the empty tube). Note that this estimate is obtained *without* changing any parameters from the case of zero fill factor. Therefore, the test demonstrates that the pressure estimation technique is insensitive to changes in the acoustic load.

Active Damping of the Acoustic Resonances

The use of the pressure estimate technique enables feedback control of the acoustic resonances without the need for additional sensors. A set of experiments is performed to demonstrate the effectiveness of the pressure estimation technique for active damping of the cavity resonances. The tube with a 70% fill factor is chosen as the testbed for the active damping tests because the mechanical-to-acoustic coupling is stronger than the case of zero fill factor. The increased mechanical-to-acoustic coupling enables the demonstration of multiple-mode control with second-order feedback compensators.

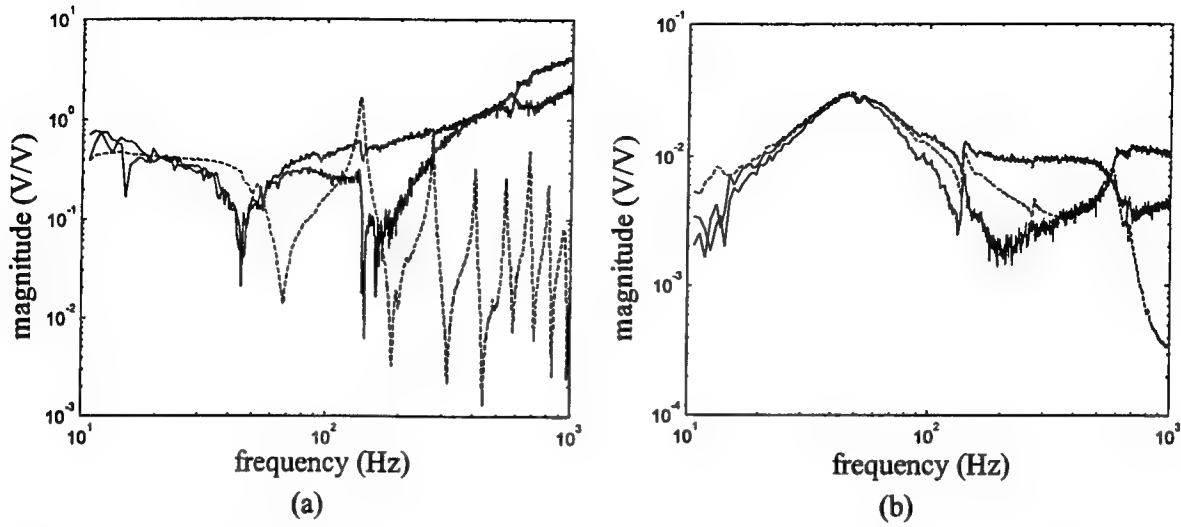


Figure 8: Effect of parameter variation on the estimated frequency responses (solid): (a) effect of varying \hat{L} on the pressure estimate and, (b) effect of varying \hat{L} on the pressure estimate. The measured pressure and velocity are shown as dashed lines.

The feedback compensator is a second-order filter of the form

$$K(s) = g_f \frac{\omega_f^2}{s^2 + 2\zeta_f \omega_f s + \omega_f^2} \quad (37)$$

where the parameters g_f , ζ_f , and ω_f are the gain, damping ratio, and natural frequency of the control filter. This type of compensator is referred to as a positive position feedback filter and has been used previously in the control of resonant systems [18, 19]. The second-order filter is implemented using the same digital signal processor used to implement the pressure and velocity estimator.

The feedback compensator is designed using the measured transfer function between the input voltage v_{in} and the pressure estimate \hat{p} . The gain, natural frequency, and damping ratio are chosen to make the gain of the series combination of $K(s)$ and $G(s) = \hat{p}/v_{in}$ greater than one in the frequency range of the first two acoustic resonances (roughly 100 to 140 Hz) while maintaining low-frequency stability near the speaker resonance (approximately 50 Hz).

The results of the feedback control implementation for the 70% fill factor are shown in Figure 10. The transfer functions illustrate the effect of active damping on the acoustic response at both ends of the tube. Increasing the gain of the compensator adds damping to the acoustic resonances at 100 Hz and 132 Hz. The damping in the 100 Hz mode increases from an open-loop value of 4.4% to a closed-loop value of 14.93% critical. The damping in the acoustic resonance at 132 Hz increases from 2.4% critical to 3.23% critical. The acoustic response actually becomes larger near the low-frequency resonance of the speaker due to the small phase margin in the frequency response of $K(j\omega)G(j\omega)$. The small phase margin is attributed to the phase lag associated with the mechanical resonance of the speaker.

The control results illustrate the utility of the pressure estimation technique for actively suppressing interior acoustic resonances. Over the frequency range in which the pressure estimate is valid – approximately 20 to 150 Hz for the present tests – the pressure estimate enables active suppression of the acoustic resonances with rather simple feedback compensators. The simplicity of the feedback compensation and the robustness

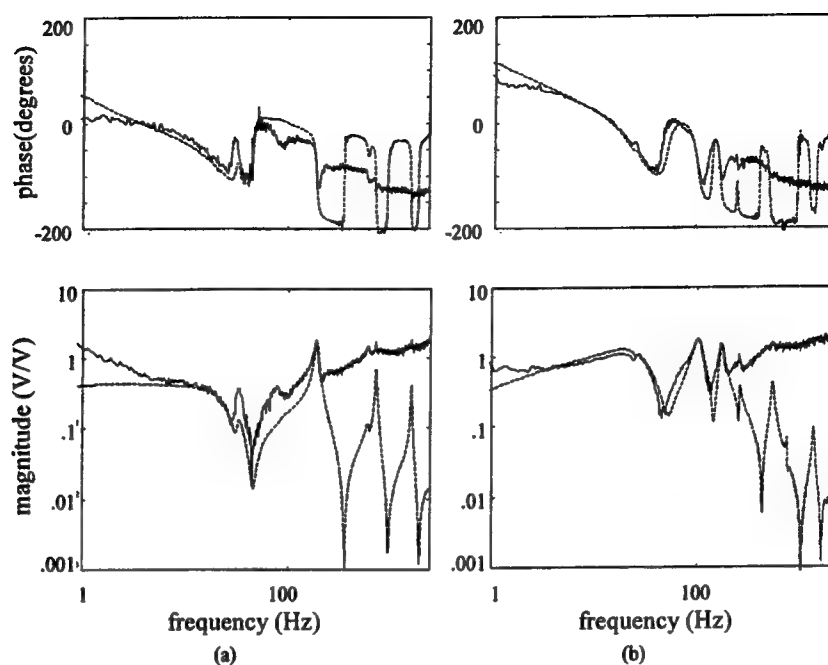


Figure 9: Pressure estimates after tuning the frequency responses (solid) and the measured pressure response (dashed): (a) zero fill factor, and, (b) 70% fill factor.

of the estimation technique is advantageous for systems in which the interior acoustic resonances are uncertain or time varying.

Acknowledgements

This work was supported by the 1998 Summer Faculty Research Program sponsored by the Air Force Office of Scientific Research. The author would like to thank his focal points, Dr. Steven Griffin and Dr. Dino Sciulli of the Dynamics Branch of the AFRL Phillips Site, for supporting this summer research. He would also like to thank the Mechanical, Industrial, and Manufacturing Engineering Department of the University of Toledo for their support during the research phase of this program.

Conclusions

A 'self-sensing' technique was developed for estimating the mechanical velocity and acoustic pressure from the measurement of speaker voltage and current. The technique was derived from an impedance model that incorporated the electrical and mechanical dynamics of the speaker in addition to the dynamics of the coupled acoustic field. The advantages of the technique is that it allows the measurement of speaker velocity and acoustic pressure without the need for additional sensors and it is insensitive to changes in the acoustic load. This makes it applicable to the control of uncertain or time-varying acoustic resonances.

The estimation technique was demonstrated on a enclosure that consisted of a closed cylindrical tube with a speaker mounted at one end. The current and voltage of the speaker were measured with an analog circuit

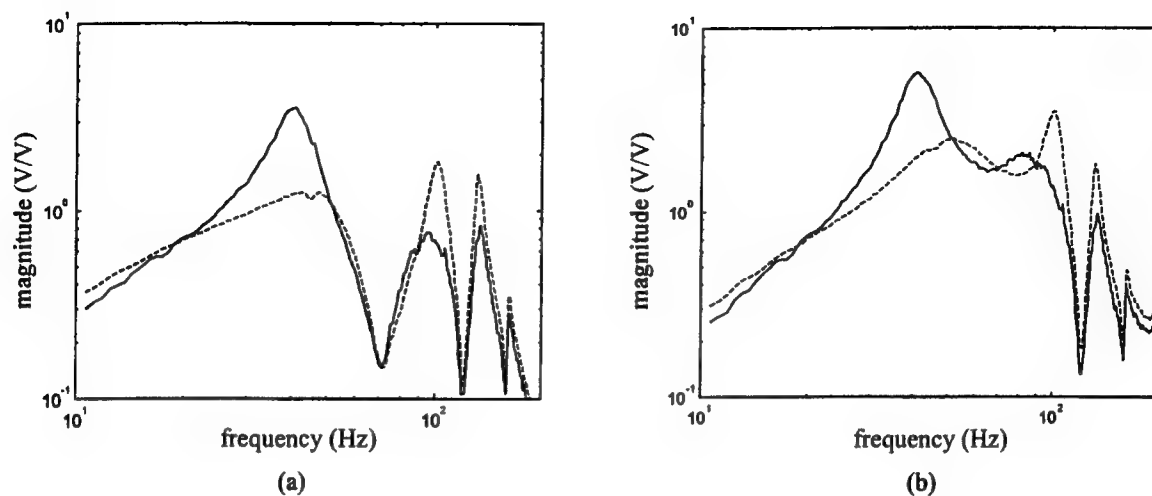


Figure 10: Open-loop (dashed) and closed-loop (solid) feedback control results utilizing the pressure estimate: (a) Microphone 1; (b) Microphone 2.

and the pressure and velocity estimator was implemented digitally. The estimates of the speaker resistance, inductance, and mechanical impedance were determined from measured transfer functions. Experimental results indicated that an accurate pressure estimate could be obtained over the frequency range 20 to 200 Hz. Errors in the pressure estimate at frequencies greater than 200 Hz were attributed to the weaker than expected mechanical-to-acoustic coupling and imperfect estimation of the electrical impedance. The insensitivity of the estimation technique to changes in the acoustic load was demonstrated by placing a cylinder inside the tube and measuring the estimated transfer function. The fill factor of the internal cylinder was approximately 70% and the first acoustic resonance dropped from 140 Hz to 100 Hz.

Active damping of the acoustic resonances was demonstrated on the tube with the 70% fill factor. A second-order feedback compensator was implemented to add active damping to the first two acoustic resonances. Experimental results demonstrated that feedback control increased the damping in the 100 Hz mode from 4.4% critical to 14.9% critical, while the damping in the second acoustic resonance increased from 2.4% critical to 3.2% critical. The small phase margin near the mechanical resonance of the speaker caused an increase in the acoustic response in the 25 to 50 Hz frequency range. Future work will concentrate on eliminating the detrimental effects of the mechanical resonance on the closed-loop response.

References

- [1] H.F. Olson and E.G. May. Electronic sound absorber. *Journal of the Acoustical Society of America*, 25:1130-1136, 1953.
- [2] W.B. Conover. Fighting noise with noise. *Noise Control*, 2:78-82, 1956.
- [3] P.A. Nelson and S.J. Elliot. *Active Control of Sound*. Academic Press, London, 1992.
- [4] S.J. Elliot, I.M. Stothers, P.A. Nelson, A.M. McDonald, D.C. Quinn, and T. Saunders. The active control of engine noise inside cars. In *Proceedings of the Inter-Noise Conference*, pages 987-990, 1988.

- [5] S.J. Elliot, P.A. Nelson, I.M. Stothers, and C.C. Boucher. In-flight experiments on the active control of propoeller-induced cabin noise. *Journal of Sound and Vibration*, 140:219-238, 1990.
- [6] P.A. Nelson and S.J. Elliot. Active minimization of acoustic fields. *Journal of Theoretical and Applied Mechanics Supplement*, 6:39-98, 1987.
- [7] P.A. Nelson, A.R.D. Curtis, S.J. Elliot, and A.J. Bullmore. The active minimization of harmonic enclosed sound fields, part i: Theory. *Journal of Sound and Vibration*, 117(1):1-13, 1987.
- [8] S.J. Elliot, P. Joseph, P.A. Nelson, and M.E. Johnson. Power output minimization and power absorption in the active control of sound. *Journal of the Acoustical Society of America*, 90(5):2501-2512, 1991.
- [9] F. Orduna-Bustamante and P.A. Nelson. An adaptive controller for active absorption of sound. *Journal of the Acoustical Society of America*, 91(5):2740-2747, 1992.
- [10] Jeffrey Dosch, Daniel J. Inman, and Ephraim Garcia. A self-sensing piezoelectric actuator for collocated control. *Journal of Intelligent Material Systems and Structures*, 3(1):166-185, 1992.
- [11] Eric H. Anderson, Nesbitt Hagood, and J.M. Goodliffe. Self-sensing piezoelectric actuation: Analysis and application to controlled structures. In *AIAA Paper Number 92-2465-CP*, 1992.
- [12] Daniel G. Cole and Robert L. Clark. Adaptive compensation of piezoelectric sensoriactuators. *Journal of Intelligent Material Systems and Structures*, 5:665-672, 1995.
- [13] Steven A. Lane and Robert L. Clark. Improving loudspeaker performance for active noise control applications. *submitted to the Journal of the Audio Engineering Society*.
- [14] Steven A. Lane and Robert L. Clark. Active control of a reverberant enclosure using an approximate constant volume velocity source. In *Proceedings of the American Control Conference*, pages 2606-2610, 1998.
- [15] Donald J. Leo, Eric H. Anderson, and Roger Glaese. Vibroacoustic modeling of a launch vehicle payload fairing for active acoustic attenuation. In *Proceedings of the Adaptive Structures Forum*, 1998.
- [16] Steven Griffin, Colin Hansen, and Ben Cazzalato. Feedback control of structurally radiated sound into enclosed spaces using structural sensing. *Submitted to the Journal of the Acoustical Society of America*.
- [17] Leo L. Beranek. *Acoustics*. McGraw-Hill Book Company, 1954.
- [18] Donald J. Leo and Daniel J. Inman. Modeling and control simulations of slewing flexible frame containing active members. *Smart Materials and Structures Journal*, 2:82-95, 1993.
- [19] Jeffrey J. Dosch, Donald J. Leo, and Daniel J. Inman. Modeling and control for vibration suppression of a flexible active structure. *Journal of Guidance, Control, and Dynamics*, 18:340-346, 1995.

EFFECT OF MATERIALS AND DESIGN VARIATIONS ON AMTEC CELL LOSSES

**M. A. K. Lodhi
Professor
Department of Physics**

**Texas Tech University
MS 1051
Lubbock, TX 79409-1051**

**Final Report for
Summer Faculty Research Program
Air Force Research Laboratory
Kirtland Air Force Base**

**Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, DC
and
Air Force Research Laboratory
Kirtland Air Force Base**

July 1998

EFFECT OF MATERIALS AND DESIGN VARIATIONS ON AMTEC CELL LOSSES

M. A. K. Lodhi

Professor

Department of Physics

Texas Tech University

Abstract

This document deals with the improvement of power output and the efficiency of the vapor-anode multi-tube AMTEC cells. In the recent past a number of programs have focused on developing AMTEC technology associated with its fabrication and design. The performance level, however, achieved to-date is hitherto below the theoretical potential of this device. In improving performance characteristics we examined the sources of power loss and proposed some solution by demonstrating with computer calculations. Two kind of losses reduce efficiency are thermal (as a result of thermal conduction and radiation of materials) and electrical (related to the ohmic resistance of material). Each of the losses can, in principle, be reduced separately by varying the current in the load. To reduce the thermal losses, the current must be increased; to reduce the electrical losses, the current in the load must be decreased. In such inversely competitive situation an optimum value is sought by applying the optimization theory. Heat losses due to radiation may be reduced by increasing the current density and by reducing the emissivity of electrodes and other surfaces. Changing the geometry (design) of the cell has proved to be helpful in improving the cell performance. As a result of this overall effort we have been able to demonstrate the improvement in the efficiency of AMTEC cell by 100%.

EFFECT OF MATERIALS AND DESIGN VARIATIONS ON AMTEC CELL LOSSES

M. A. K. Lodhi

Introduction

A power system considered for a number of potential space missions requires long life, high specific power (power/mass), that is low mass, high areal power density (power/ area), high efficiency, low cost and static in operation. To a great extent a system called Alkali Metal Thermal to Electric Converter (AMTEC) does possess latent qualities that would qualify it to be a possible candidate for space power. It can provide efficiency close to the theoretical Carnot efficiency at relatively low temperatures. AMTEC has conversion efficiency much higher than other direct thermoelectric devices. An optimized AMTEC can potentially provide a theoretical efficiency when operated between 1,000 K to 1,300 K on the hot side and between 400 K to 700 K on the condenser side. It is fuel source insensitive. It can utilize heat as input fuel from most of any source like fossil fuel, the Sun, radioisotopes, or the nuclear reactor. AMTEC, with solar energy as a heat source, is capable of being an alternative to photovoltaic-based power system for use of low earth orbit (LEO) for future NASA and Air Force missions. It is intended to be used for future NASA missions like Pluto Express (PX) and Europa early in the millennium 2000 with radioisotope decay as its heat input.

Since middle sixties a number of programs in developing the operating principle, design and technology of AMTEC have been evolving rapidly. Kummer and Weber demonstrated the conversion of heat, through the sodium cycle, into electricity by the use of beta" alumina solid electrolyte (BASE) in a patent assigned to Ford Motor Company in 1968 [1]. Several years later Weber described the operating principle of AMTEC with a liquid anode in a historical paper in 1974 [2]. It took some time for the community to recognize the potential qualities of AMTEC with regard to its uses and application for terrestrial and space power. Accordingly, there was not much activity in modeling, design and technology development of AMTEC until almost the beginning of 1990s. During that AMTEC dormant period there were few landmark papers primarily on its principle and working efficiency[3-7]. A review article on the progress of thermionic technology during a ten-year period of 1983-92 covers the slow moving development in AMTEC

- technology during that period [8]. The use of nuclear power, from the radioisotopes or nuclear reactors, for the space pointing out the potential use of AMTEC perhaps provided another motivation in looking for its application [9-11]. A flurry of AMTEC activities began with a great vigor around the world at the beginning of 1990s. NASA Lewis Research Center (LeRC) has an interest in developing thermal energy storage (TES) for the solar dynamic ground test demonstration (SDGTD) program. The research and development effort at Jet Propulsion Lab (JPL) included studies, which address both overall device construction and investigation into the AMTEC components [12-25]. Advanced Modular Power Systems (AMPS) have been focusing on designing and manufacturing AMTEC cells for Pluto Express mission [26-27]. Air Force Research Lab (AFRL) has an interest in developing an electrical power system for the payload in low earth orbit (LEO) for a duration of five years and in testing the performance of AMTEC for PLUTO express [28-30]. Orbital Science Corporation has been conducting a series of studies of radioisotopes power system based on general-purpose heat source (GPHS) for potential deep space missions [9,31-33]. AMTEC has been found compatible with direct conversion from the projected SP-10 nuclear space power reactor [10-11]. Recently University of New Mexico has been engaged in modeling and analysis of AMTEC performance and evaluation [34]. In Japan Kyushu University in collaboration with Electrotechnical Laboratory the interest in AMTEC analysis has evolved [35-36]. Creare Inc. has primarily concentrated on evaporator component of AMTEC to enable it operate efficiently [37-39].

While these programs resolved a number of key technological issues successfully associated with the design and fabrication of AMTEC, the performance level achieved hitherto is still below the theoretical potential of this device. In tracking down the source of power loss during the test cell performance of various AMTEC designs the major losses are found to be radiative and conductive. Since the cell is tested in vacuum there is no convective loss. Most of the heat escapes through the cell wall. There are radiative losses due to condenser and along the artery as well. In order to minimize these losses we propose to use some different materials which should be compatible with the operating conditions of the cell and capable of reducing these losses. Some variation in the geometry of the cell configuration is also utilized in minimizing the loss of power. We select a test AMTEC cell designated by PX-3A, which has been in

operation at US Air Force Research Laboratory since July 9, 1997 and compare its theoretical performance as a function of the variation in the material and geometrical parameters.

Principle and Working of AMTEC Cell

A typical AMTEC cell is shown in Fig. 1. An AMTEC cell is a static device for the conversion of heat to electricity performing two distinct cycles: 1) conversion of heat to mechanical energy via a sodium – based (or any suitable alkali metal) heat engine and 2) conversion of mechanical energy to electrical energy by utilizing the special properties of the BASE material. It has the potential for reliable long life operation and high conversion efficiency. The general principles governing the operation of an AMTEC cell have been given quite elaborately in the early stages of its development [1-5]. The essential thermodynamic aspects of the sodium-based engine have been described in those references. Accordingly a brief outline will be presented here for the self-consistency of this work.

AMTEC is a relatively new type of device, based on the principle of a sodium concentration cell, conceived in late sixties [1-4]. A closed vessel is divided into a high temperature, high-pressure region in contact with a heat source and a low temperature, low pressure region in contact with a heat sink. A barrier of a BASE sheet whose ionic conductivity is much larger than its electronic conductivity separates these regions. The BASE is coated with a porous metal electrode (anode) which covers the low-pressure surface of the BASE. A closed container is partially filled with a small quantity (typically <10g) of liquid sodium as the working fluid. Sodium ions disperse through the BASE in response to the pressure differential (gradient of Gibbs free energy). Electrical leads are connected with the porous anode and with the high temperature liquid (vapor) sodium, which acts as the cathode. When the circuit is closed electrons flow to the porous anode surface through the load, producing electrical work. A return line and electromagnetic pump or a wick circulates the sodium from the cold zone (condenser) to the hot zone (heater) of AMTEC. It is limited to 1350 K because of sodium interaction with the BASE. The lower temperature region is limited to a minimum of 500 K by the need of maintaining the sodium in the liquid state (but it would be operating around 625 K for high efficiency). As the condenser temperature increases above 700K the efficiency decreases [3].

The main component of an AMTEC cell is the BASE tube which essentially conducts sodium ions much more rapidly than what it does for neutral sodium atoms or negative particles like electrons. A sodium pressure difference across a thin BASE tube drives sodium ions from the high-pressure side to the low-pressure side. Thus positively charged sodium ions concentrate on the low-pressure side while negatively charged particles, electrons remain concentrated on the high-pressure side (inside the BASE tube), resulting an electrical field gradient across the BASE tube. This gradient balances the sodium pressure differential thus preventing the further flow of sodium ions. This electrical potential difference is utilized to drive an electric current through a load by applying some appropriate electrodes at the two ends of the tube. The unusual properties of the BASE provide a method of converting the mechanical energy, represented by the pressure differential, into electrical energy. Or in other words the BASE converts a chemical potential difference into an electrical potential difference. As the electrons, after going through the load, reunite with sodium ions they form neutral sodium atoms in the vapor state on the low-pressure side. The sodium vapor travels to the condenser where it condenses into liquid state. The sodium liquid is pressurized through a wick or electromagnetic pump and is returned to the high-pressure side of the BASE. In that way the thermal-to-mechanical-to-electrical conversion process is completed. The efficiency of this final conversion is governed by variety of irreversible kinetic and transport processes occurring at the electrode interfaces, within the BASE material, internal impedance, and thermal conduction and radiation losses [12-14]. A number of efforts are underway to develop practical and high efficiency cells [18, 40-45]. AMPS, in collaboration with AFRL, has been manufacturing and testing for improving efficiency of PX series of AMTEC cells [26-30]. Specifications of some of the latest multi-tube AMTEC cells are given in Table I. The longest in operation among them is PX-3A, which we have modified theoretically for improving efficiency by varying parameters, which govern its design and material. We first investigate theoretically the optimum conditions for the efficiency with respect to the geometrical parameters and material properties and then prepare a computer model of the cell to predict its performance. As a basis of the computer program we are using the APEAM CODE [34].

AMTEC Theory and Analysis

The transport of sodium through an AMTEC cell is a complex phenomenological process. Its exact detailed analysis is very difficult and is further complicated as it requires the simultaneous solution of thermal, fluid flow, and electrical analysis. Those analyses are interdependent, with each of the three analyses requiring the results of the other two. This interdependence is between a number of axially varying distribution functions. (Had it been between single-valued variables ; it could have been solved by a set of simultaneous algebraic equations.) Specifically , solving for the cell's temperature distribution requires the knowledge of the axial variation of the sodium flux through the BASE tube and of the electrical output power density profile over the tube length. Solution of the axial pressure variation of the low-pressure sodium requires knowledge of the cell's temperature distribution and the BASE tubes' current density variation. Similarly, solving for the axial variation of the current density and of the inter-electrode voltage requires prior solution of the axial variation of the BASE tube temperature and internal-to-external pressure ratio. Those interdependent distribution functions require solutions of coupled differential and integral equations by a more sophisticated procedure. Schock *et al* [32] generated a thermal analysis model for multi-tube AMTEC cell by appropriately modifying the ITAS and SINDA codes [46,47].

In the absence of any externally applied electric field the internal electric field arises due to the non-uniform distribution of charged species in the BASE tube. The electrochemical potential, e responsible to transport sodium ions is given by [48]

$$e = M + F\phi,$$

where M is the chemical potential, ϕ is the electrical potential acting on sodium ions and F is the Faraday constant equal to 96,500C/mole. The force F on the ion is thus :

$$\mathbf{F} = - 1/N \text{ grad } e = - 1/N \text{ grad } (M + F \phi),$$

Where N is Avagadro's number. If \mathbf{F} is directed along the thickness of the BASE, denoted by x ; the flux is a function of the electric field $d\phi/dx$. The sodium flow \mathbf{J} is given by:

$$\mathbf{J} = - CB/N (\text{grade}) = CB/N (dM/dx + Fd\phi/dx) \mathbf{n},$$

where C is the concentration, B the mobility and \mathbf{n} is the unit vector along the normal. If the BASE is brought to equilibrium ; the electrochemical potential on side I, denoted by e_I will be the same as e_{II} on side II of the BASE wall so that $e_I = e_{II}$ implies for $\mathbf{grad} e$ to vanish. Thus,

$$(M_I - M_{II}) \mathbf{n} = -F \mathbf{grad} \phi.$$

This is general form of Nernst equation [5]. Energy conversion devices have few equilibria and are typically open systems unlike the classical thermodynamics which is restricted to reversible and closed systems. Onsager's treatment of irreversible processes, such as diffusion, can be applied to AMTEC operation to deal with its irreversibility of the process and openness of the system. One can write, in principle, the effective emf, V as a function of cell voltage in open-circuit, V_{oc} and electrode polarization overpotential, $V_{op} (= \zeta^a - \zeta^c)$ [2] from Nernst equation [5]. The open-circuit voltage and charge-exchange current density are related with the cell temperature and pressure [19]. We can thus write for the net power output as:

$$P = VI,$$

where

$$V = V_{oc} - IR_{int} - V_{op},$$

$$V_{oc} = RT_B/F \ln P_a/P_c,$$

I is the net current in the circuit and R_{int} is the total internal resistance of the cell including the contact resistance of the electrodes, sheet resistance in the plane of electrodes, resistance of the current collectors, bus wires and conductor leads to the load, the charge exchange polarization losses at the BASE-electrodes interfaces, and the BASE ionic resistance, R_B given by: $R_B = \rho_B t_B$.

R is the gas constant = 8.314 J/mol.K. T_B is the temperature of the BASE tube, P_a and P_c are the pressures at the anode and cathode sides respectively, ρ_B and t_B are electrical resistivity, the expression for which has been developed by Steinbruck [49], and thickness of BASE tube. R_{int} will be written here as the sum of the terms in the same order as stated above:

$$R_{int} = R_{contact} + R_{sheet} + R_{collector} + R_{bus} + R_{op} + R_{BASE}$$

Heat Losses

There are many sources of heat loss due to conduction, Joule heating and radiation. The conduction heat loss, Q_{cd} and Joule heating, Q_J are given by:

$$Q_{cd} = I\lambda^{1/2} [T_B \cos\delta - T_{cond}] / \sin\delta$$

And

$$Q_J = I\lambda^{1/2} [T_B + T_{cond}] \times [1 - \cos\delta] / \sin\delta,$$

where $\delta = I\lambda^{1/2} L/kA_l$, I is the total cell current, λ is Franz-Lorentz's number $= 2.45 \times 10^{-8} \text{ W}\Omega/\text{K}^2$, L is the length of a conductor lead, k is the thermal conductivity of a lead, A_l is the cross sectional area of a connecting lead, T_B and T_{cond} are temperatures of BASE tube and the condenser respectively. The radiation heat dissipated from the BASE tube to both the condenser and the cell wall is given by:

$$Q_{rad} = \sigma/Z [T_B^4 - T_{cond}^4],$$

where σ is Stefan Boltzman constant and Z is given by [50]:

$$Z = 1/\epsilon_B + 1/\epsilon_{cond} - 1 + N[2/\epsilon_s - 1],$$

where ϵ_B , ϵ_{cond} , ϵ_s are emissivities of BASE, condenser and radiation shield respectively and N is the number of radiation shields between the BASE tube and cell wall.

There are two major losses of power inherent to thermal cycle of the working fluid in the cell. One is in cooling of the fluid from temperature T_B to T_{cond} in which case the heat loss, Q_{th} is given by:

$$Q_{th} = c_p (T_B - T_{cond}) IM,$$

where c_p is the heat capacity at constant pressure, and for liquid sodium it is $1,285 \text{ J/kg.K}$ [51]. The other major loss of power is in condensing the vapor metal to liquid. The latent heat of vaporization, h_L for sodium liquid is determined by the well-known Clausius-Clapeyron equation [52]:

$$h_L = 5383.2 IM \ln 10 = 4.480 \text{ kJ/kg}$$

Conversion Efficiency

The overall conversion efficiency of AMTEC cell is given by [5,52]:

$$\eta = VI / \{ I[V + (h_L + Q_{th}) / F] + Q_{loss} \},$$

where Q_{loss} is defined as:

$$Q_{loss} = Q_{cd} + Q_{rad}.$$

In order to maximize the efficiency we can write its reciprocal as:

$$1/\eta = 1 + \{ I [h_L + c_p (T_B - T_{cond})] / F + Q_{loss} \} / I(RT_B/F \ln P_a/P_c - IR_{int} - V_{op}).$$

From the theory of optimization, for the maximum efficiency, the behavior of the parameters involved is determined as follows:

- 1) High-value parameters: T_B, P_a
- 2) Low-value parameters: $h_L, c_p, (T_B - T_{cond}), Q_{loss}, P_c, I, R_{int}, V_{op}$

This implies that:

- The choice of working fluid be such that its latent heat of vaporization and specific heat must be low
- The emissivities of materials used for BASE tube, outer condenser surface, outer artery surface, inner cell wall and radiation shields must be low for minimizing radiation losses
- Conductive and Joule losses through the leads are small compared with other losses. They are, however, competitive with each other. The former requires the thermal conductivity and cross sectional area of the leads to be small whereas the later one favors their larger values. There are some optimum values of these parameters, which correspond to maximize the efficiency.
- The output power maximizes when the load resistance is equal to the internal resistance of the cell. This implies that the load resistance should match the internal resistance for the best results.
- To make good contact of BASE tubes with the heat source the studs' areas should be maximized
- There corresponds an optimum number of BASE tubes to a minimum radiation loss

- To increase the number of radiation shields to an optimum value helps reducing the radiation losses.

We then proceeded with these constraints as input data for the computer program in order to get the resulting output power and efficiency.

Results

From the optimization analysis in the previous section we take two approaches, one by considering the properties of materials and the other from the cell geometry, which essentially change the design of the cell. Later on we combined the two approaches to look for the best results. The results are given in figures 2-4.

Conclusion

We have been able to reduce the losses almost on all counts thereby increasing the output power and thus the efficiency of AMTEC. The efficiency has improved by 100 % when we incorporate the variation in the cell design (geometrical parameters) and cell materials. For further reducing the heat losses and thus improving the efficiency we make the following recommendations:

- The interior cell wall be coated with some appropriate material with high reflective coefficient and low emissivity
- The cell wall could be made of double layers with air in between
- The artery's outer surface should be coated with the same material used for the cell wall coating
- Condenser may be coated with sodium also
- The power output maximizes at a certain value of the current in the circuit, and when the load resistance matches the internal resistance of cell. For the best results the cell must be operated under those conditions.
- The condenser's temperature should be kept optimum for the maximum output power and efficiency.

References

1. Kummer, J.T. and N. Weber, U.S. Patent 3,458,356 (1968), Assigned to Ford Motor Co
2. Weber, N. "A Thermoelectric Device Based on Beta-Alumina Solid Electrolyte", *Energy Conversion* **14**, 1 (1974)
3. Hunt, T.K., N. Weber, and T. Cole, "Research on the Sodium Heat Engine", *Proc. 13th Intersoc Energy Conversion Engineering Conf.*, SAE, Warrendale, PA (1978) p 201
4. Hunt, T. K., N. Weber, and T. Cole, "High Efficiency Thermoelectric Conversion with Beta-Alumina Solid Electrolyte: The Sodium Heat Engine", *Fast Ionic Transport in Solids*, ed. J.B. Bates, G.C. Ferrington (N Holland Co, 1981) p 263
5. Cole, T. "Thermoelectric Energy Conversion with Solid Electrolyte," *Science*, **221**, 915 (1983)
6. Ewell, R. and J. Mondt, "Static Conversion Systems", *Space Nuclear Power Sys.*, Ed. M. S. El-Genk and M.D. II Hoover (Orbit Book Co. Malabar, FL, 1985) pp 385-391
7. Bankston, C. P., T. Cole, S. K. Khanna and A. P. Thakoor, "Alkali Metal Thermodynamic Conversion (AMTEC) for Space Nuclear Power Systems", *Space Nuclear Power Sys.* Ed. M. S. El-Genk and M.D. Hoover, (Orbit Co., Malabar, FL, 1985) pp 398-402
8. Dahlberg, R. C., *et al*, "Review of Thermionic Technology 1983 to 1992," "A Critical Rev of Space Nuclear Power and Propulsion 1984 =1992", Ed. M. S. El-Genk (AIP Press, NY 1994) pp 121-161
9. Shock, A. "Design Evolution and Verification of the General Purpose Heat Source," *Proc 15th Intersocieties Energy Conversion Engineering Conf.* Vol. 2 (AIAA, NY 1980) p 1032
10. Birden, D. and F.A. Angelo, "Space Reactor Past, Present and Future", *Proc 18th Intersocieties Energy Conversion Engineering Conf.*, Vol (AIChE, N.Y., 1983) p 61
11. Brat, E.J. and G.O. Fitzpatrick, "Direct Conversion Nuclear Reactor Space Power Sys." Report No. AFWALTR-82-2073, Vol 1
12. Williams, R. A. *et al* " High power density performance of WPt and WRn electrodes in the alkali metal thermoelectric converter," *J. Electrochem. Soc.* **136**, 893 (1989)
13. Williams, R. A. *et al* " Kinetics and transport at AMTEC electrodes, I. The interfacial impedance model," *J. Electrochem. Soc.* **137**, 1709 (1990)

14. Williams, R. A. *et al* "Kinetics and transport at AMTEC electrodes, II. Temperature dependence of the interfacial impedance of Na(g)/porous Mo/ Na- β " alumina," *ibid* **137**, 1716 (1990)
15. Rayan, M. A. *et al* " Directly deposited current collecting grids for AMTEC electrodes." Proc. 26th Intersociety Energy Conversion Engineering Conf. **5**, 463-468 (1991)
16. Rayan, M. A. *et al* " Vapor exposure test cell for thin film electrodes on solid electrolytes," JPL New Technology Report , 18620-8166 (1991)
17. Rayan, M. A. *et al* " AMTEC electrode morphology changes as studied by electrochemical impedance spectroscopy and other techniques," Proc. Symp. On High Temperature Electrode Materials, Electrochem. Soc. **91-6**, 115 (1991)
18. Vinning, C. B., R.M. Williams, M.L. Underwood, M.A. Ryan, and J.W. Suiter, "Reversible Thermodynamic Cycle for AMTEC Power Conversion", Proc 27th Intersocieties Energy Conversion Engineering Conf., Society of Automobile Engineers, 1992 **3**: , pp 123-128
19. Underwood, M. A. *et al* "An AMTEC vapor-vapor, series connected cell," Proc. 9th Symp. On Space Nuclear Power Systems, ed. M. S. El-Ghenk and M. D. Hoover, American Inst. Phys. **3**, 1331 (1992)
20. Underwood, M. A. *et al* " Efficiency of an AMTEC recirculating test cell," Proc. 27th IECEC, **3**, 197 (1992)
21. Underwood, M. A. *et al* " A five-volt AMTEC multicell," Proc. 28th IECEC, **1**, 855 (1993)
22. Rayan, M. A. *et al* " Vapor exposure test cell for testing thin film electrodes," NASA Tech. Brief80, (1993)
23. Rayan, M. A. *et al* " Development in AMTEC devices, components and performance," CONF940101 (AIP Press NY) 1495 (1994)
24. Rayan, M. A. *et al* " Lifetimes of thin film AMTEC electrodes," Proc. 29th IECEC, **2**, 877 (1994)
25. Rayan, M. A. *et al* " Lifetime modeling of Ti N electrodes for AMTE C CELLS," Proc. Space Tech. And Application Intl. Forum, ED M. S. El-Ghenk, (AIP Press NY) 1607 (1998)

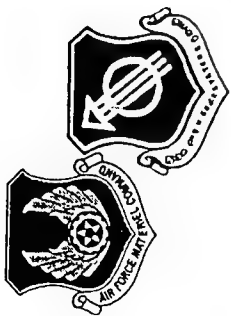
26. Sievers, R. K. *et al*, “ ,” Proc. 10th SSNPP, CONF. (AIP Press, NY) 319 (1993)
27. Sievers, R. K. *et al*, “ ,” Proc. STAIF (AIP Press, NY) 1479 (1998)
28. Schuller, M. J. *et al*, “Proc. 31st IECEC, IEEE, 877 (1996)
29. Merrill, J. *et al*. Proc. 32nd IECEC, American Chemical Soc. , 1184 (1996)
30. Merrill, J. *et al*. Proc., STAIF, (AIP Press NY) 1613, (1998)
31. Schock, A. H. “Design evolution and verification of the general purpose of heat source,” Proc. 15th Intersoc. Energy Conversion Engineering Conf. , 1032 (1980)
32. Schock, A. H. *et al*, “Design, analysis, and fabrication procedure of AMTEC cell, test assembly, and radioisotope power system for outer-planet missions,” Proc. 48th Intl. Astronomical Congress, 1 (1997) and references therein
33. Schock, A. H. *et al*, “OSC design and validated analysis of AMTEC generator for Europa orbiter mission,” Preprint for presentation in 33rd Intesoc. Energy Conversion Engineering Conf. 1998, and references therein
34. Tournier, J-M. and M. S. El-Ghenk, “ AMTEC performance and evaluation analysis model (APEAM): comparison with test results of PX-4C,PX-3A cells,” Proc. 15th STAIF (AIP Press NY) 1576 (1998), and references therein
35. Kato, A., Nakata, and K. Tsuchida, “Characteristics of ceramic electrode for AMTEC,” Proc. 27th IECEC, American Chemical Soc. 3 (1992), Paper No. 929006
36. Kato, A., Nakata, and K. Tsuchida, “Characteristics of ceramic electrode for AMTEC – II, structure of electrode and power density,” Proc. 28th IECEC, American Chemical Soc. 809 (1993), Paper No.93027
37. Crowley, C. J. and M. G. Izenson, “Condensation of sodium on a micromachined surface from AMTEC,” Proc. 10th Symp. On Space Nuclear Power and Propulsion,” (AIP Press, NY) 897 (1993)
38. Izenson, M. G. and C. J Crowley,”Micromachined evaporator for AMTEC cells,” *ibid* I.829
39. Izenson, M. G.,”Micromachined evaporator for AMTEC cells,” IEEE, 2226 (1996)

40. Sayer, M. *et al*, J. Applied Physics, **67**, 832 (1990)
41. Underwood, M. L. *et al*, J. Propulsion and Power, (1993)
42. Underwood, M. L. *et al*, "Recent Advances in AMTEC Recirculating Test Cell Performance", Conf 930103 (AIP, 1993) pp 885-890
43. Schöck, A. *et al*, "Effect of design variations of AMTEC cell efficiency. And of operating parameters on performance of OSCcell design," Proc. 32nd Interdisciplinary Energy Conversion Conf. #97529 (1997)
44. Lodhi, M.A.K. Michael Schuller, and Paul Housgen, "Mathematical modifying for a Thermionic AMTEC Cascade System," Proc space tech and applications international forum 1996, ed. M. S. El-Genk, (AIP Press, NY) pp1285-1294
45. El-Ghenk, M. S. and J-M. Tournier, "Recent advances in vapor-anode, multi-tube, alkali metal thermal-to-electric conversion cells for space power," Proc. 5th ESPS-98 (to be presented)
46. Noravian, H. "ITAS. Integrated Thermal Analysis System." Proc. 26th Intl Conf. On Environmental Systems.#961376 (1996)
47. Gaski, J. "SINDA (System Improved Numerical Differencing Analyzer), vers. 1.315 from Network Analysis Associate, Fountain Valley, CA. 1987
48. Straus A. M. and S. W. Peterson, "Irreversible thermodynamics of AMTEC devices," Proc. STAIF , 1571, 1998
49. Steinbruck, M. *et al*, "Investigation of beta -alumina solid electrolytes for application in AMTEC cells," Proc. 28th IECEC I.1, 799, 1993
50. Siegel, R. and J. R. Howell, "Thermal radiation heat transfer," (Hemisphere publishing Corp., New York, 1981) pp 283-292
51. Tournier, J-M., M. S. El-Ghenk, M. Schuller and P. Hausgen, "An analytical model for liquid-anode and vapor-anode AMTEC converters," Proc. 14th STAIF-97, 397, 1997
52. Saad, M. A., "Thermodynamics for Engineers," (Prentice-Hall, Inc. New Jersey, 1966) pp 243-245

Table 1. Design Parameters of PX-Series Cells Tested at AFRL.

Cell design	PX-1B	PX-1C	PX-2A	PX-2C	PX-4C	PX-5A	PX-5B	PX-3G	PX-3A	PX-3C
Diameter (mm)	38.1	38.1	38.1	38.1	38.1	38.1	38.1	38.1	31.75	31.75
Cell height (mm)	101.6	101.6	88.9	88.9	101.6	101.6	101.6	101.6	101.6	101.6
Evaporator	DC	DC	DC	DC	DC	DC	SC	SC	DC	flat
Elevation (mm)	12.7	12.7	3.81	3.81	3.81	3.81	3.81	3.81	5.18	7.62
Standoff thick. (mm)	0.71	0.71	0.71	1.5	0.71	0.71	0.71	0.71	0.71	0.71
Standoff material	SS	SS	SS	SS	SS	SS	SS	SS	SS	Ni
Standoff rings (mm)	-	-	-	-	2.5	2.5	1.25	1.25	1.1	-
Rings material	-	-	-	-	SS	SS	Ni	Ni	Ni	-
Conduct. stud (mm ²)	-	-	100	100	100	100	100	100	38	OSC
Stud material	-	-	SS	SS	SS	SS	SS	SS	SS	Ni
# of BASE tubes	7	7	7	7	6	6	6	6	5	5
Tube length (mm)	27	27	40	40	40	40	40	40	32	40
Electrode/tube (mm ²)	300	300	600	600	600	600	600	600	600	600
Tube braze material	TiCuNi	TiCuNi	TiCuNi	TiCuNi	TiNi	TiNi	TiNi	TiNi	TiNi	TiNi
Current collector	100-mesh Cu	100-mesh Cu	100-mesh Cu	100-mesh Cu	60-mesh Mo	60-mesh Mo	60-mesh Mo	60-mesh Mo	60-mesh Mo	60-mesh Mo
Feedthrough braze	TiCuAg	TiCuAg	TiCuNi	TiCuNi	TiCuNi	TiCuNi	TiCuNi	TiCuNi	TiCuNi	TiCuNi
Radiation shield Material	-	Circular SS	chevrons SS	Circular SS	Circular SS	Circular SS	Circular Ta	Circular SS	Circular SS	(Rh coating)
Condenser type	mesh pad	mesh pad	mesh pad	mesh pad	mesh pad	Creare	Creare	Creare	Creare	Creare
Hot side	SS	SS	SS	SS	SS	SS	SS	SS	SS	Ni
Cell wall	SS	SS	SS	SS	SS	SS	SS	SS	SS	Haynes-25
Initial test date	9/1/96	11/10/96	11/25/96	12/24/96	3/1/97	5/1/97	8/14/97	11/4/97	7/9/97	11/26/97
Operation (hrs)	240	240	96	600	1800	672	3,729	1,590	4,500	2,040

DC = Deep Cone; SC = Shallow Cone; OSC = Orbital Sciences Corporation advanced stud/plenum design.



Multi-BASE Tube AMTEC Cell

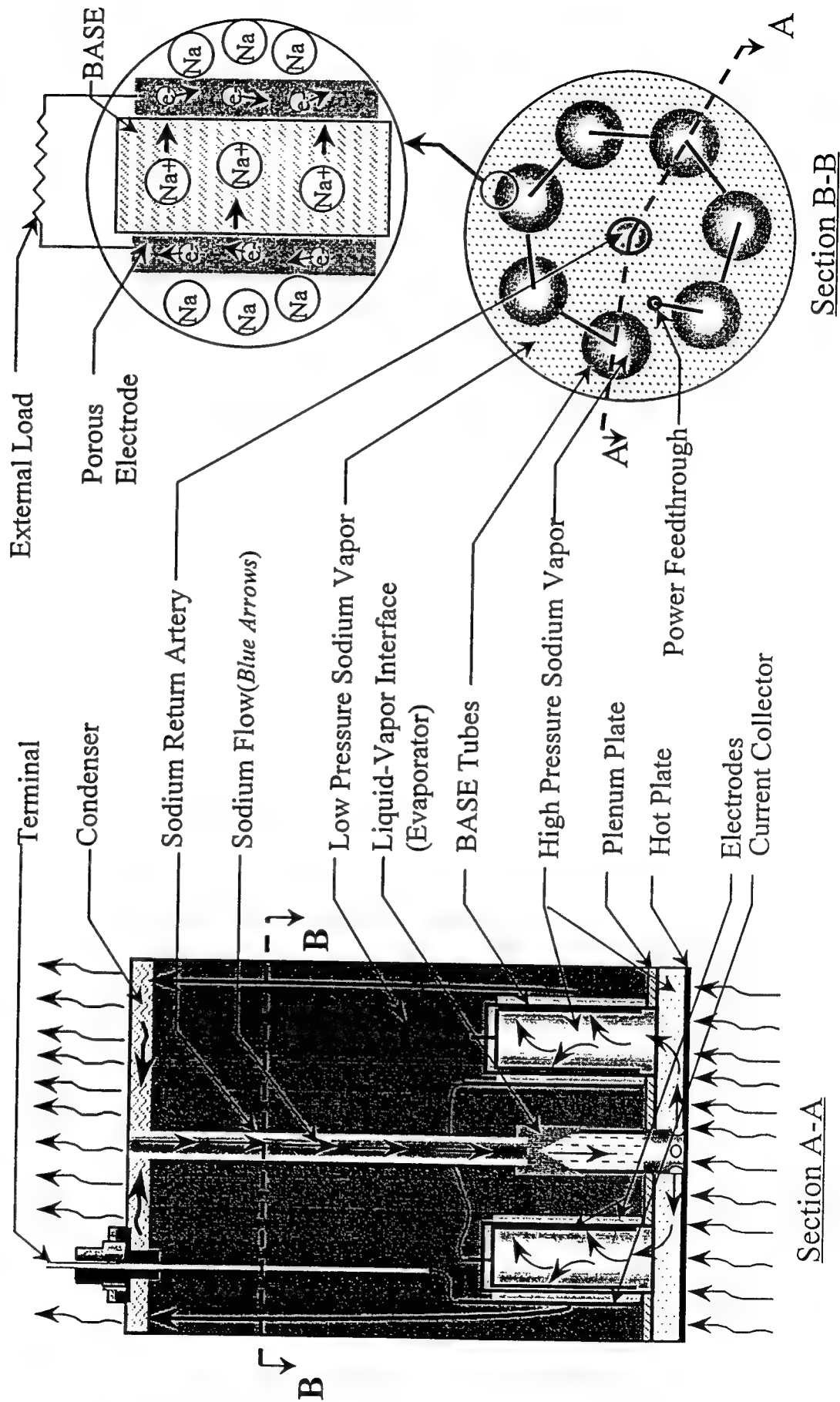


Figure 1

Cell Power versus External Load for PX3A

$T_{hot}=1173\text{ K}$, $T_{cond}=623\text{ K}$

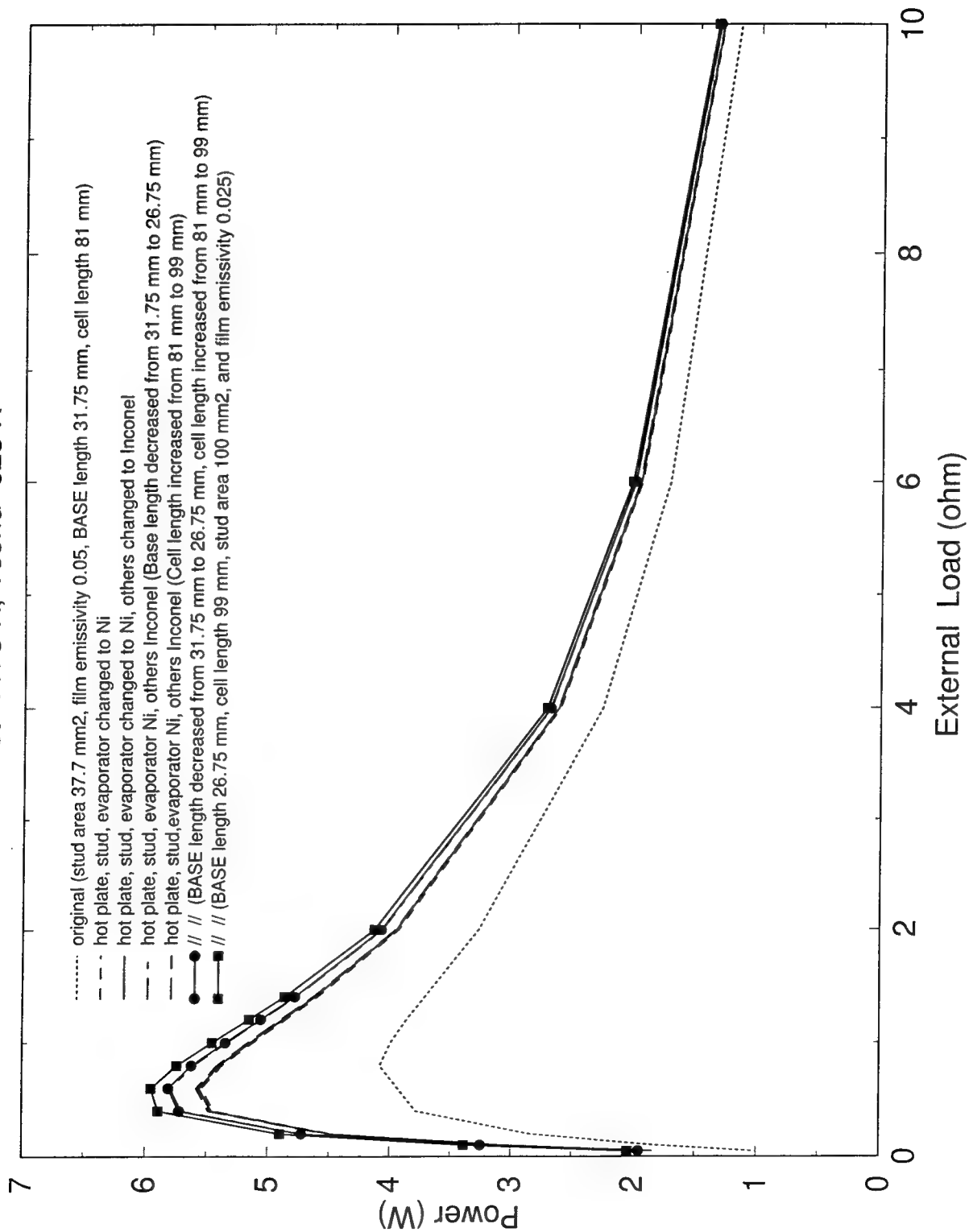


Figure 2

Cell Efficiency for PX3A

$T_{hot}=1173\text{ K}$, $T_{cond}=623\text{ K}$

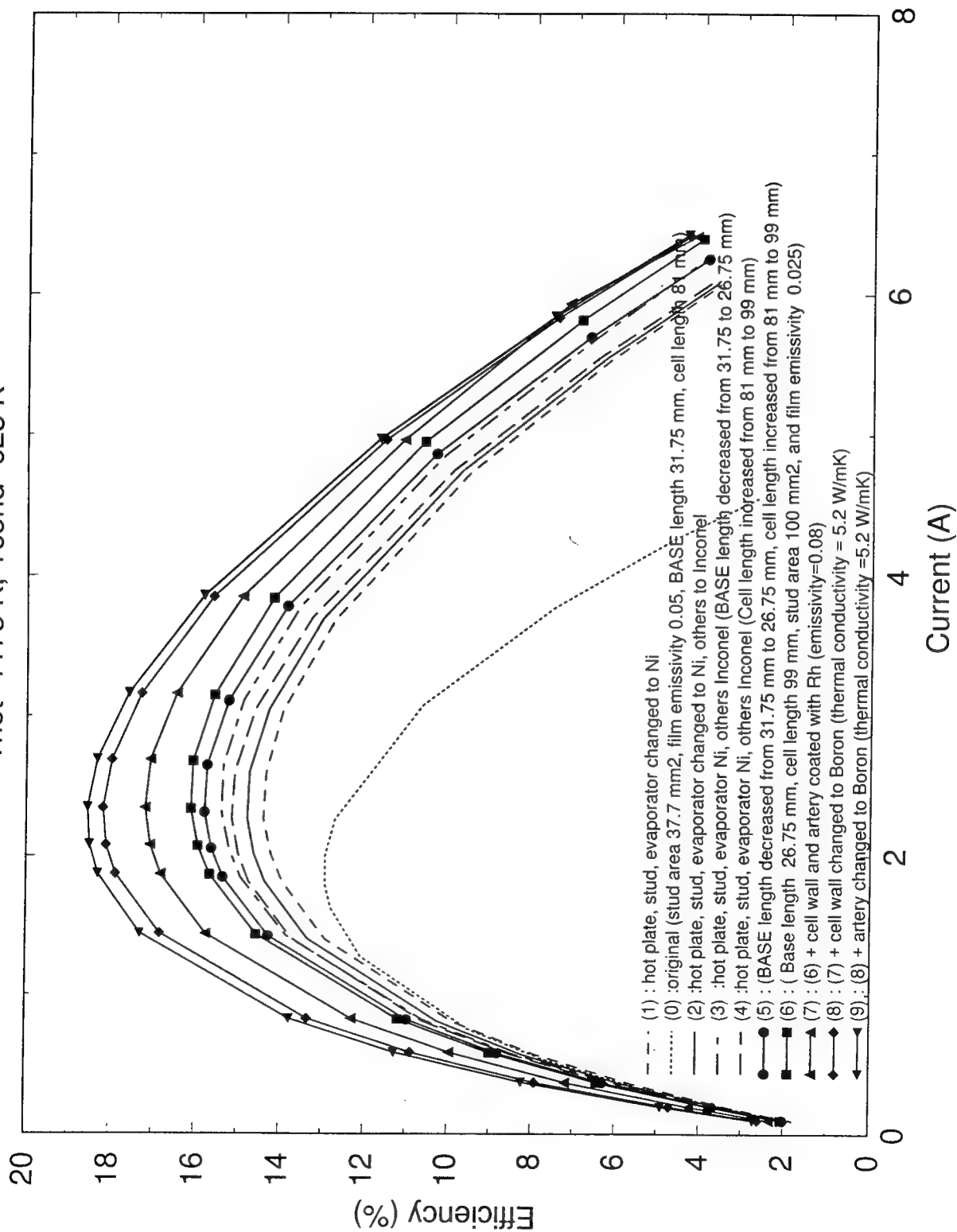


Figure 3

Cell Efficiency for PX3A

$T_{hot}=1173\text{ K}$, $T_{cond}=623\text{ K}$

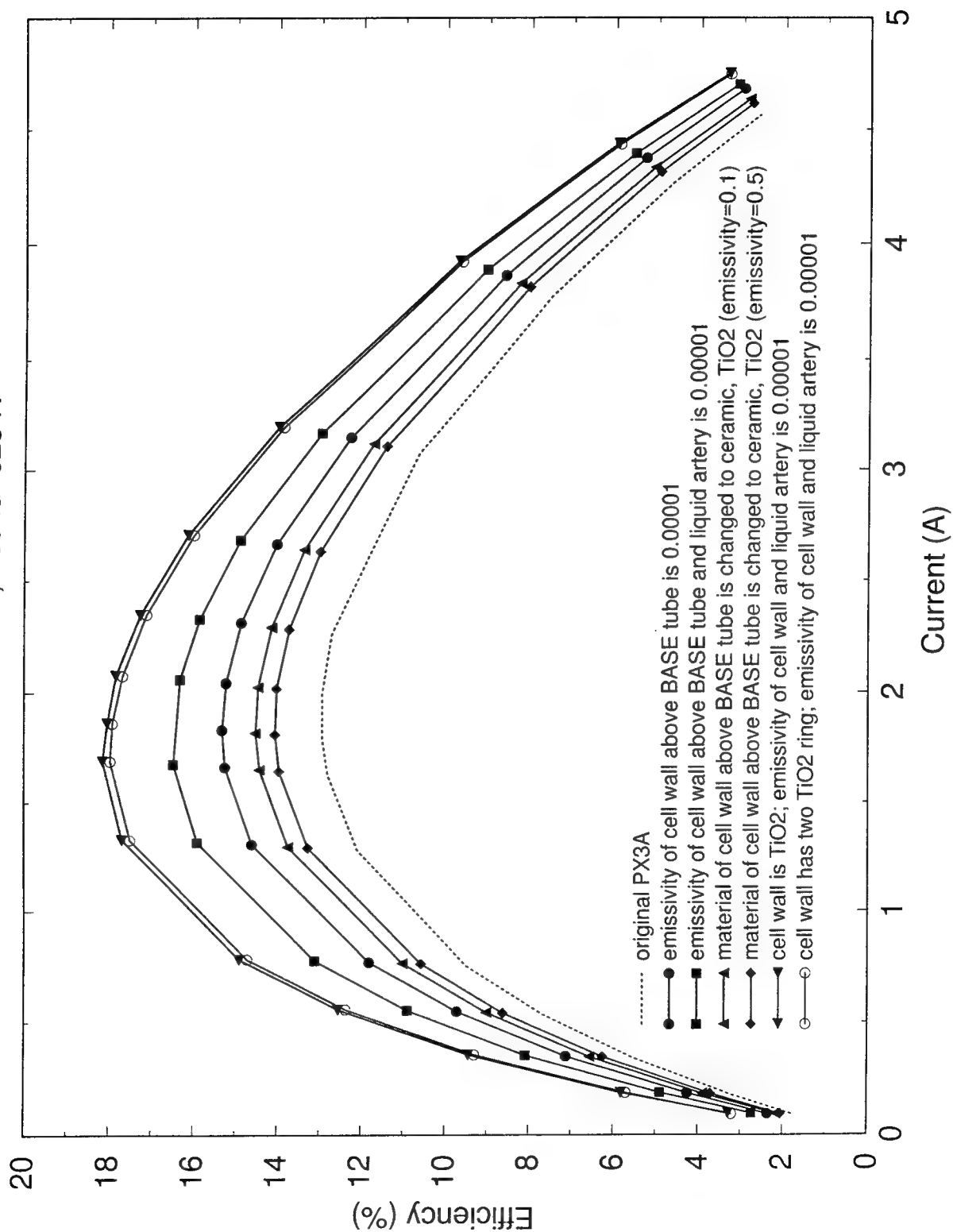


Figure 4

A splitting technique for the anelastic equations in atmospheric physics

**John P. McHugh
Associate Professor
Department of Mechanical Engineering**

**Department of Mechanical Engineering
University of New Hampshire
Durham, NH 03824**

**Final Report for:
Summer Faculty Research Program
Hanscom Research site**

**Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC and**

Hanscom Research Site

October, 1998

A SPLITTING TECHNIQUE FOR THE ANELASTIC EQUATIONS IN ATMOSPHERIC PHYSICS

John P. McHugh
Associate Professor
Department of Mechanical Engineering

Abstract

Gravity waves propagating in the earth's atmosphere are studied numerically. The anelastic approximation to the equations for an adiabatic atmosphere are treated. The temporal discretization is a splitting algorithm. The buoyancy terms in the vertical momentum equation and the energy equation treated are treated by an analytic manipulation to isolate equations for vertical velocity and potential temperature. The pressure is determined not with Poisson's equation, but a similar elliptic operator chosen to conform to the continuity equation for the anelastic equations. Preliminary results show wave saturation and a variation in wavelength when the base temperature profile has a sudden change in Brunt-Vaisala frequency.

A splitting technique for the anelastic equations in atmospheric physics

John P. McHugh

1 Introduction

Atmospheric turbulence is a phenomenon which is of primary interest to Air Force operations. Turbulence interferes with flight operations, satellite communications, and weapons systems operation. Turbulence is also pervasive in the atmosphere, and cannot be avoided. Recent measurements of turbulent intensity has shown that the elevation near the tropopause experiences excessively high levels of turbulence. The turbulence seems to form in thin layers, on the order of hundreds of meters thick, but with very large horizontal extent, on the order of tens of kilometers. The term 'blini' has been associated with these patches of turbulence. The blini's are believed to be caused by atmospheric gravity waves, which propagate with an upwards component of velocity, reach a critical elevation and 'break', and finally disintegrate into turbulence. The ability to predict the occurrence of blinis, and other high levels of turbulence, is greatly desired.

A numerical method for the study of the propagation and breaking of gravity waves is considered. Much previous work has considered the numerical solution of gravity waves in the atmosphere, as outlined below. However this work has focussed on constant temperature cases where there is no tropopause. The present is concerned with cases which do not have constant temperature.

The standard model used in the study of atmospheric flow is the adiabatic atmosphere, where the gas is assumed to obey the perfect gas law, all processes are assumed to be reversible, and heat transfer between fluid particles is neglected. Often the assumption of zero heat transfer is modified by allowing turbulent exchange of energy. The numerical solution of the resulting fully compressible equations has been found to accurately model many important phenomena, see for example the recent gravity wave results of Fritts [8]. However, there is a major drawback to the fully compressible equations, namely that sound waves are generated. These sound waves do not significantly effect other phenomena, such as convection, waves, or gravity currents, but they must be numerically resolved in time or

the numerical methods will fail to converge. The presence of sound waves therefore imposes an unfortunately stringent restriction on the time step, and puts many flows of interest outside the reach of direct numerical simulations of the fully compressible equations.

Two other models, which are based on the fully compressible adiabatic atmosphere model, have also been used to study atmospheric dynamics with computers. These models are the Boussinesq equations and the anelastic equations. The Boussinesq equations have been known for over one hundred years, and their properties are pretty well known. The basic assumption behind them is that the temperature does not change very rapidly. Holton [7] discusses the compressible form of the Boussinesq equations.

The anelastic equations were developed in the sixties [1],[3],[2]. The basic idea is to use the time scale of an oscillation as a parameter, and separate the terms which are responsible for sound waves from the other dynamics. The original anelastic equations have a flaw in that the linear solution does not match the linear solution of the fully compressible equations. It was later shown that the pressure term was incorrect. Bacmeister and Schoeberl [4] provide a concise derivation of the corrected anelastic equations.

There has been much previous work studying atmospheric gravity waves using numerical models. Recent work by Fritts [8] considered a direct simulation of fully three-dimensional motion. The governing equations were the fully compressible perfect gas equations, and the Reynolds number was small, many orders of magnitude smaller than the atmospheric Reynolds number. In spite of the low Reynolds number, the results of Fritts, et al., are insightful and provide much needed information about the mechanics of an internal breaking waves and the resulting turbulence. However the base state temperature profile was constant temperature, hence there was no tropopause. Furthermore, the choice of the fully compressible equations, while convenient and potentially more accurate, imposes the severe restriction on the time step mentioned above. Walterscheid and Schubert [5] also used the fully compressible equations, using only two-dimensions, and provide basic information on the breaking of gravity waves before turbulence becomes prominent. Walterscheid and Schubert also assumed a uniform temperature for the base state.

Bacmeister and Schoeberl [4] studied gravity waves propagating in two-dimensions using the anelastic equations. The anelastic equations have the promise of being capable

of treating larger scale flows where the scale height varies significantly in the domain of interest. Bacmeister and Schoeberl also assume a constant temperature base state for their results.

2 Governing equations

The anelastic equations for a compressible atmosphere are

$$\frac{Du}{Dt} = -\frac{\partial p^*}{\partial x} + \frac{1}{R_e} \nabla^2 u, \quad (1)$$

$$\frac{Dv}{Dt} = -\frac{\partial p^*}{\partial y} + \frac{1}{R_e} \nabla^2 v, \quad (2)$$

$$\frac{Dw}{Dt} = -\frac{\partial p^*}{\partial z} + \frac{1}{F_r^2} \frac{\theta}{\bar{\theta}} + \frac{1}{R_e} \nabla^2 w, \quad (3)$$

$$\frac{D\theta}{Dt} + w \frac{\partial \bar{\theta}}{\partial z} = \frac{1}{R_e P_r} \nabla^2 \theta, \quad (4)$$

$$\frac{\partial \bar{\rho} u}{\partial x} + \frac{\partial \bar{\rho} v}{\partial y} + \frac{\partial \bar{\rho} w}{\partial z} = 0, \quad (5)$$

where

$$p^* = c_p \bar{\theta} \left(\frac{p}{p_0} \right)^{\frac{R}{c_p}}, \quad (6)$$

u, v, w are the velocity components, x, y, z are the components of position, θ is the potential temperature, $\bar{\rho}$ and $\bar{\theta}$ are the basic state density and potential temperature, c_p is the specific heat at constant pressure, R is the gas constant, p is the pressure, p_0 is a constant, and the Reynolds, Prandtl, and Froude numbers are

$$R_e = \frac{UL}{\nu}, \quad (7)$$

$$P_r = \frac{\nu}{\kappa}, \quad (8)$$

$$F_r^2 = \frac{U^2}{gL}. \quad (9)$$

Equations (1-3) are the momentum equations, (4) is the energy equation, and (5) is the continuity equation.

The length scales of atmospheric flows are generally large; thunderstorms, gravity waves, and other phenomena are measured in kilometers. The effect of the large scale is to make molecular viscous diffusion and molecular heat diffusion negligible. Energy is dissipated in such flows by a transfer of energy to turbulence, and the turbulence turns the energy into internal energy through molecular viscosity. But the direct effect of molecular viscosity on gravity waves, and other large scale phenomena, is very small. If there is no resident turbulence in the flow, or if turbulence is not being generated, the flow is essentially inviscid. We first consider a splitting algorithm for the inviscid flow case.

3 Temporal integration for inviscid flow

Consider the inviscid equations. The time integration is accomplished using a splitting technique, similar to Karniadakis [6], and many others. Each equation is integrated in time from t_n to t_{n+1} . Using equation (??) as an example,

$$\int_{t_n}^{t_{n+1}} \frac{\partial w}{\partial t} dt = - \int_{t_n}^{t_{n+1}} \frac{\partial p^*}{\partial z} dt - \int_{t_n}^{t_{n+1}} \left[u \frac{\partial w}{\partial x} + v \frac{\partial w}{\partial y} + w \frac{\partial w}{\partial z} \right] dt + \int_{t_n}^{t_{n+1}} \frac{1}{F_r^2 \bar{\theta}} \theta dt. \quad (10)$$

The integral involving pressure can be simplified by defining an average pressure, \tilde{p} :

$$\tilde{p} = \int_{t_n}^{t_{n+1}} p^* dt. \quad (11)$$

This step is convenient, although it is not necessary.

Each of the terms on the right-hand-side of (10) may be evaluated separately using the splitting idea. Experience has shown that the linear term must be evaluated using an implicit scheme, such as the Adams-Moulton method, for stability. The nonlinear term is evaluated with an explicit scheme, such as the Adams-Bashforth method, for efficiency:

$$\int_{t_n}^{t_{n+1}} \frac{1}{F_r^2 \bar{\theta}} \theta dt = \sum_{q=0} \alpha_q \frac{1}{F_r^2 \bar{\theta}} \theta^{n+1-q} = \sum_{q=0} \alpha_q W_w^{n+1-q}, \quad (12)$$

$$\int_{t_n}^{t_{n+1}} \left[u \frac{\partial w}{\partial x} + v \frac{\partial w}{\partial y} + w \frac{\partial w}{\partial z} \right] dt = \sum_{q=0} \beta_q \left[u \frac{\partial w}{\partial x} + v \frac{\partial w}{\partial y} + w \frac{\partial w}{\partial z} \right]^{n-q} = \sum_{q=0} \beta_q N_w^{n-q}, \quad (13)$$

where W_w is the wave operator, N_w is the nonlinear convective operator, and α_q and β_q are constants pertaining the particular choice of integration scheme. Note that the subscript w signifies that N_w (or W_w) is the component of a vector function, \mathbf{N} (or \mathbf{W}), from the vertical momentum equation. The resulting equation for w is

$$\frac{w^{n+1} - w^n}{\Delta t} = -\frac{\partial \tilde{p}}{\partial z} - \sum_{q=0} \beta_q N_w^{n-q} + \sum_{q=0} \alpha_q W_w^{n+1-q}. \quad (14)$$

It is convenient to consider velocity and potential temperature to be parts of a single solution vector,

$$\Phi = \begin{pmatrix} u \\ v \\ w \\ \theta \end{pmatrix}. \quad (15)$$

The discreet equations may then be written as

$$\Phi^{n+1} = \Phi^n + \mathbf{P}^{n+1} + \sum_{q=0} \alpha_q \mathbf{W}^{n+1-q} - \sum_{q=0} \beta_q \mathbf{N}^{n-q}, \quad (16)$$

where \mathbf{P} are the pressure terms,

$$\mathbf{P} = \begin{pmatrix} \frac{\partial \tilde{p}}{\partial x} \\ \frac{\partial \tilde{p}}{\partial y} \\ \frac{\partial \tilde{p}}{\partial z} \\ 0 \end{pmatrix}, \quad (17)$$

\mathbf{W} indicates the so-called wave terms,

$$\mathbf{W} = \begin{pmatrix} 0 \\ 0 \\ \frac{1}{F_r^2 \theta} \theta \\ -\frac{\partial \theta}{\partial z} w \end{pmatrix}, \quad (18)$$

\mathbf{N} are the nonlinear terms,

$$\mathbf{N} = \begin{pmatrix} u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} + w \frac{\partial u}{\partial z} \\ u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} + w \frac{\partial v}{\partial z} \\ u \frac{\partial w}{\partial x} + v \frac{\partial w}{\partial y} + w \frac{\partial w}{\partial z} \\ u \frac{\partial \theta}{\partial x} + v \frac{\partial \theta}{\partial y} + w \frac{\partial \theta}{\partial z} \end{pmatrix}, \quad (19)$$

and the boldface means a vector quantity.

We would like to obtain the solution by splitting the calculation into two substeps; the wave and nonlinear terms, and the pressure terms. The steps may be written as

$$\frac{\hat{\Phi} - \Phi^n}{\Delta t} = \sum_{q=0} \alpha_q \mathbf{W}^{n+1-q} - \sum_{q=0} \beta_q \mathbf{N}^{n-q}, \quad (20)$$

$$\frac{\Phi^{n+1} - \hat{\Phi}}{\Delta t} = -\mathbf{P}^{n+1}. \quad (21)$$

This is similar to Karniadakis, who studied incompressible flow and included viscous terms.

4 The wave terms

The wave terms in (20) cause some difficulty when treated implicitly, as described here. Ignoring the nonlinear terms for the moment, only the vertical momentum equation and the energy equation have a nontrivial contribution in (20). These two equations are

$$\frac{\hat{w} - w^n}{\Delta t} = \frac{1}{F_r^2 \bar{\theta}} \sum_{q=0} \alpha_q \theta^{n+1-q}, \quad (22)$$

$$\frac{\hat{\theta} - \theta^n}{\Delta t} = -\frac{\partial \bar{\theta}}{\partial z} \sum_{q=0} \alpha_q w^{n+1-q}. \quad (23)$$

The difficulty with a direct application of these equations is that w^{n+1} and θ^{n+1} are not yet known at the first step of the splitting algorithm. Since the wave terms have a simple form, the discrete equations may be manipulated to eliminate the problem terms. However, the manipulation must occur before splitting.

The discrete equations for w and θ from (16) are

$$w^{n+1} = w^n + \Delta t \left[-\frac{\partial \tilde{p}}{\partial z} - \sum_{q=0} \beta_q N_w^{n-q} + \sum_{q=0} \alpha_q W_w^{n+1-q} \right], \quad (24)$$

$$\theta^{n+1} = \theta^n - \Delta t \left[\sum_{q=0} \beta_q N_\theta^{n-q} + \sum_{q=0} \alpha_q W_\theta^{n+1-q} \right]. \quad (25)$$

Extract the first term from the summation in the wave terms to get

$$w^{n+1} = w^n + \frac{\Delta t \alpha_0}{F_r^2 \bar{\theta}} \theta^{n+1} + \Delta t \left[-\frac{\partial \tilde{p}}{\partial z} - \sum_{q=0} \beta_q N_w^{n-q} + \sum_{q=1} \alpha_q W_w^{n+1-q} \right], \quad (26)$$

$$\theta^{n+1} = \theta^n - \Delta t \alpha_0 \bar{\theta}_z w^{n+1} - \Delta t \left[\sum_{q=0} \beta_q N_\theta^{n-q} + \sum_{q=1} \alpha_q W_\theta^{n+1-q} \right]. \quad (27)$$

Note that both equations contain w^{n+1} and θ^{n+1} . Eliminate w^{n+1} from (51) and θ^{n+1} from (50). The result is

$$\begin{aligned} [1 + \Delta t^2 B_v^2 \alpha_0^2] w^{n+1} &= w^n + \Delta t \left[-\frac{\partial \tilde{p}}{\partial z} - \sum_{q=0} \beta_q N_w^{n-q} + \sum_{q=1} \alpha_q W_w^{n+1-q} \right], \\ &+ \frac{\Delta t \alpha_0}{F_r^2 \bar{\theta}} \left[\theta^n - \Delta t \left(\sum_{q=0} \beta_q N_\theta^{n-q} + \sum_{q=1} \alpha_q W_\theta^{n+1-q} \right) \right] \end{aligned} \quad (28)$$

$$\begin{aligned} [1 + \Delta t^2 B_v^2 \alpha_0^2] \theta^{n+1} &= \theta^n - \Delta t \left[\sum_{q=0} \beta_q N_\theta^{n-q} + \sum_{q=1} \alpha_q W_\theta^{n+1-q} \right] \\ &- \Delta t \alpha_0 \bar{\theta}_z \left[w^n + \Delta t \left(-\frac{\partial \tilde{p}}{\partial z} - \sum_{q=0} \beta_q N_w^{n-q} + \sum_{q=1} \alpha_q W_w^{n+1-q} \right) \right] \end{aligned} \quad (29)$$

where B_v is the Brunt-Vaisala frequency,

$$B_v^2 = \frac{1}{F_r^2} \frac{\bar{\theta}_z}{\bar{\theta}}. \quad (30)$$

Splitting for these two equations is achieved with the following two steps:

$$\begin{aligned} \frac{\hat{w} - w^n}{\Delta t} = & \left[-\sum_{q=0} \beta_q N_w^{n-q} + \sum_{q=1} \alpha_q W_w^{n+1-q} \right] \\ & + \frac{\alpha_0}{F_r^2 \bar{\theta}} \left[\theta^n - \Delta t \left(\sum_{q=0} \beta_q N_\theta^{n-q} + \sum_{q=1} \alpha_q W_\theta^{n+1-q} \right) \right], \end{aligned} \quad (31)$$

$$\begin{aligned} \frac{\hat{\theta} - \theta^n}{\Delta t} = & - \left[\sum_{q=0} \beta_q N_\theta^{n-q} + \sum_{q=1} \alpha_q W_\theta^{n+1-q} \right] \\ & - \alpha_0 \bar{\theta}_z \left[w^n + \Delta t \left(-\sum_{q=0} \beta_q N_w^{n-q} + \sum_{q=1} \alpha_q W_w^{n+1-q} \right) \right], \end{aligned} \quad (32)$$

$$\frac{[1 + \Delta t^2 B_v^2 \alpha_0^2] w^{n+1} - \hat{w}}{\Delta t} = -\frac{\partial \tilde{p}}{\partial z}, \quad (33)$$

$$\frac{[1 + \Delta t^2 B_v^2 \alpha_0^2] \theta^{n+1} - \hat{\theta}}{\Delta t} = \Delta t \alpha_0 \bar{\theta}_z \frac{\partial \tilde{p}}{\partial z}. \quad (34)$$

Note that (58) and (59) are basically the same operator. Some insight can be obtained by studying the original analytic equations. Dropping all terms except the wave terms and the time derivative in (3) and (4) gives

$$\frac{\partial w}{\partial t} = \frac{1}{F_r^2} \frac{\theta}{\bar{\theta}}, \quad (35)$$

$$\frac{\partial w}{\partial t} + w \frac{\partial \bar{\theta}}{\partial z} = 0. \quad (36)$$

Eliminating θ from (35) and w from (36) results in

$$\frac{\partial^2 w}{\partial t^2} + B_v^2 w = 0, \quad (37)$$

$$\frac{\partial^2 \theta}{\partial t^2} + B_v^2 \theta = 0. \quad (38)$$

These equations are merely the second order wave equation. Note that both w and θ obey the same equation. The discrete equation given by (52) would match a discrete version

of (37) for simple difference schemes, however (52) is the higher order equation which is consistent with the original differential equations, while a discrete form of (37) may not be.

5 Pressure

The pressure solution is obtained with an elliptic operator, while simultaneously enforcing the continuity equation. The numerical solution of the incompressible Navier-Stokes equations normally means forming the Poisson equation for pressure,

$$\nabla^2 p^{n+1} = \frac{1}{\Delta t} [\nabla \cdot \hat{\mathbf{u}} - \nabla \cdot \mathbf{u}^{n+1}] \quad (39)$$

The continuity equation for incompressible flow,

$$\nabla \cdot \mathbf{u}^{n+1} = 0, \quad (40)$$

is then used to eliminate $\nabla \cdot \mathbf{u}^{n+1}$ on the right hand side. This is fortuitous, since \mathbf{u}^{n+1} is usually unknown at this stage.

The continuity equation for the anelastic equations is

$$\nabla \cdot \bar{\rho} \mathbf{u}^{n+1} = 0, \quad (41)$$

which means that the Poisson equation for pressure is

$$\nabla^2 \bar{p}^{n+1} = \frac{1}{\Delta t} \left[\nabla \cdot \hat{\mathbf{u}} + \frac{\bar{\rho}_z}{\bar{\rho}} w^{n+1} \right]. \quad (42)$$

This presents a problem, since w^{n+1} has not yet been determined.

The quantity $\nabla \cdot \bar{\rho} \nabla \bar{p}$ may be formed instead of $\nabla^2 \bar{p}$, which would allow for the elimination of w^{n+1} from the equation, except for the coefficient $[1 + \Delta t^2 B_v^2 \alpha_0^2]$, which causes difficulty. However, a similar operator can be formed, resulting in

$$\frac{\partial}{\partial x} \left(\bar{\rho} \frac{\partial \bar{p}}{\partial x} \right) + \frac{\partial}{\partial y} \left(\bar{\rho} \frac{\partial \bar{p}}{\partial y} \right) + \frac{\partial}{\partial z} \left(\frac{\bar{\rho}}{[1 + \Delta t^2 B_v^2 \alpha_0^2]} \frac{\partial \bar{p}}{\partial z} \right) =$$

$$\frac{1}{\Delta t} \left[\frac{\partial}{\partial x} (\bar{\rho} \hat{u}) + \frac{\partial}{\partial y} (\bar{\rho} \hat{v}) + \frac{\partial}{\partial z} \left(\frac{\bar{\rho} \hat{w}}{[1 + \Delta t^2 B_v^2 \alpha_0^2]} \right) \right] \quad (43)$$

$$\frac{1}{\Delta t} [-\nabla \cdot \bar{\rho} \mathbf{u}^{n+1}]. \quad (44)$$

We then choose continuity to be enforced at the leading time step, and use (41) to eliminate the troublesome term on the right-hand-side.

The normal derivative of pressure is imposed as pressure boundary conditions, and is obtained from (16). For example, a bounday with a normal in the z direction has

$$\frac{\partial \bar{p}^{n+1}}{\partial z} = w^n - w^{n+1} + \sum_{q=0} \alpha_q W_w^{n+1-q} - \sum_{q=0} \beta_q N_w^{n-q}. \quad (45)$$

In the absence of the viscous terms, a direct application of (45) does not present the usual problems associated with pressure boundary conditions.

6 Temporal integration for viscous flow

The discreet equations may now be written as

$$\Phi^{n+1} = \Phi^n + \mathbf{P}^{n+1} + \sum_{q=0} \alpha_q \mathbf{W}^{n+1-q} - \sum_{q=0} \beta_q \mathbf{N}^{n-q} + \sum_{q=0} \alpha_q \mathbf{L}^{n+1-q}, \quad (46)$$

where \mathbf{L} are the viscous dissipation terms,

$$\mathbf{L} = \begin{pmatrix} \frac{1}{Re} \nabla^2 u \\ \frac{1}{Re} \nabla^2 v \\ \frac{1}{Re} \nabla^2 w \\ \frac{1}{Re Pr} \nabla^2 \theta \end{pmatrix}. \quad (47)$$

The discreet equations for w and θ are now

$$w^{n+1} = w^n + \Delta t \left[-\frac{\partial \bar{p}}{\partial z} - \sum_{q=0} \beta_q N_w^{n-q} + \sum_{q=0} \alpha_q W_w^{n+1-q} + \sum_{q=0} \alpha_q L_w^{n+1-q} \right], \quad (48)$$

$$\theta^{n+1} = \theta^n - \Delta t \left[\sum_{q=0} \beta_q N_\theta^{n-q} + \sum_{q=0} \alpha_q W_\theta^{n+1-q} + \sum_{q=0} \alpha_q L_\theta^{n+1-q} \right]. \quad (49)$$

Proceeding as before, extract the first term from the summation in the wave terms (but not the diffusion terms) to get

$$w^{n+1} = w^n + \frac{\Delta t \alpha_0}{F_r^2 \bar{\theta}} \theta^{n+1} + \Delta t \left[-\frac{\partial \tilde{p}}{\partial z} - \sum_{q=0} \beta_q N_w^{n-q} + \sum_{q=1} \alpha_q W_w^{n+1-q} + \sum_{q=1} \alpha_q L_w^{n+1-q} \right], \quad (50)$$

$$\theta^{n+1} = \theta^n - \Delta t \alpha_0 \bar{\theta}_z w^{n+1} - \Delta t \left[\sum_{q=0} \beta_q N_\theta^{n-q} + \sum_{q=1} \alpha_q W_\theta^{n+1-q} + \sum_{q=1} \alpha_q L_\theta^{n+1-q} \right]. \quad (51)$$

Eliminate w^{n+1} from (51) and θ^{n+1} from (50) (except for the diffusion terms). The result is

$$\begin{aligned} [1 + \Delta t^2 B_v^2 \alpha_0^2] w^{n+1} = \\ w^n + \Delta t \left[-\frac{\partial \tilde{p}}{\partial z} - \sum_{q=0} \beta_q N_w^{n-q} + \sum_{q=1} \alpha_q W_w^{n+1-q} + \sum_{q=0} \alpha_q L_w^{n+1-q} \right] \\ + \frac{\Delta t \alpha_0}{F_r^2 \bar{\theta}} \left[\theta^n - \Delta t \left(\sum_{q=0} \beta_q N_\theta^{n-q} + \sum_{q=1} \alpha_q W_\theta^{n+1-q} + \sum_{q=0} \alpha_q L_\theta^{n+1-q} \right) \right], \end{aligned} \quad (52)$$

$$\begin{aligned} [1 + \Delta t^2 B_v^2 \alpha_0^2] \theta^{n+1} = \\ \theta^n - \Delta t \left[\sum_{q=0} \beta_q N_\theta^{n-q} + \sum_{q=1} \alpha_q W_\theta^{n+1-q} + \sum_{q=0} \alpha_q L_\theta^{n+1-q} \right] \\ - \Delta t \alpha_0 \bar{\theta}_z \left[w^n + \Delta t \left(-\frac{\partial \tilde{p}}{\partial z} - \sum_{q=0} \beta_q N_w^{n-q} + \sum_{q=1} \alpha_q W_w^{n+1-q} + \sum_{q=0} \alpha_q L_w^{n+1-q} \right) \right]. \end{aligned} \quad (53)$$

Splitting for these two equations is achieved with the following three steps:

$$\begin{aligned} \frac{\hat{w} - w^n}{\Delta t} = & \left[-\sum_{q=0} \beta_q N_w^{n-q} + \sum_{q=0} \alpha_q W_w^{n+1-q} \right] \\ & + \frac{\alpha_0}{F_r^2 \bar{\theta}} \left[\theta^n - \Delta t \left(\sum_{q=0} \beta_q N_\theta^{n-q} + \sum_{q=1} \alpha_q W_\theta^{n+1-q} \right) \right], \end{aligned} \quad (54)$$

$$\begin{aligned} \frac{\hat{\theta} - \theta^n}{\Delta t} = & - \left[\sum_{q=0} \beta_q N_{\theta}^{n-q} + \sum_{q=1} \alpha_q W_{\theta}^{n+1-q} \right] \\ & - \alpha_0 \bar{\theta}_z \left[w^n + \Delta t \left(- \sum_{q=0} \beta_q N_w^{n-q} + \sum_{q=1} \alpha_q W_w^{n+1-q} \right) \right], \end{aligned} \quad (55)$$

$$\frac{\hat{w} - \hat{w}}{\Delta t} = - \frac{\partial \tilde{p}}{\partial z}, \quad (56)$$

$$\frac{\hat{\theta} - \hat{\theta}}{\Delta t} = \Delta t \alpha_0 \bar{\theta}_z \frac{\partial \tilde{p}}{\partial z}, \quad (57)$$

$$\frac{[1 + \Delta t^2 B_v^2 \alpha_0^2] w^{n+1} - \hat{w}}{\Delta t} = \sum_{q=0} \alpha_q L_w^{n+1-q} + \frac{\Delta t \alpha_0}{F_r^2 \bar{\theta}} \sum_{q=0} \alpha_q L_{\theta}^{n+1-q}, \quad (58)$$

$$\frac{[1 + \Delta t^2 B_v^2 \alpha_0^2] \theta^{n+1} - \hat{\theta}}{\Delta t} = \sum_{q=0} \alpha_q L_{\theta}^{n+1-q} - \Delta t \alpha_0 \bar{\theta}_z \sum_{q=0} \alpha_q L_w^{n+1-q}. \quad (59)$$

The last terms on the right-hand-side of (58) and (59) result again in the issue of separating the w calculation from the θ calculation. If these last two terms are retained, then w and θ cannot be separated, and we must contend with a matrix of double the rank. Another alternative is to neglect these last terms, on the basis of the coefficient, Δt^2 . Since Δt is small, then its square is negligible. This assumption simplifies the calculation significantly.

7 Pressure revisited

Proceeding as before, the elliptic equation for pressure is

$$\frac{\partial}{\partial x} \left(\bar{\rho} \frac{\partial \tilde{p}}{\partial x} \right) + \frac{\partial}{\partial y} \left(\bar{\rho} \frac{\partial \tilde{p}}{\partial y} \right) + \frac{\partial}{\partial z} \left(\frac{\bar{\rho}}{[1 + \Delta t^2 B_v^2 \alpha_0^2]} \frac{\partial \tilde{p}}{\partial z} \right) = \quad (60)$$

$$\begin{aligned} & \frac{1}{\Delta t} \left[\frac{\partial}{\partial x} (\bar{\rho} u) + \frac{\partial}{\partial y} (\bar{\rho} v) + \frac{\partial}{\partial z} \left(\frac{\bar{\rho} w}{[1 + \Delta t^2 B_v^2 \alpha_0^2]} \right) \right] \\ & - \frac{1}{\Delta t} \left[\nabla \cdot \bar{\rho} \mathbf{u}^{n+1} \right] - \frac{1}{\Delta t} \left[\nabla \cdot \bar{\rho} \sum_{q=0} \alpha_q \nabla^2 \mathbf{u}^{n+1-q} \right]. \end{aligned} \quad (61)$$

Continuity may be used again to eliminate one of the troublesome terms on the right-hand-side. The other troublesome term (viscous dissipation) has been found in the past to be negligible, and is usually neglected here.

The pressure boundary conditions set the normal derivative of pressure using (46). As before, a boundary with a normal in the z direction has

$$\frac{\partial \tilde{p}^{n+1}}{\partial z} = w^n - w^{n+1} + \sum_{q=0} \alpha_q W_w^{n+1-q} - \sum_{q=0} \beta_q N_w^{n-q} + \sum_{q=0} \alpha_q L_w^{n+1-q}. \quad (62)$$

Again, there is a difficulty with a direct application of (62). The linear viscous terms contain the velocity at the leading time step, which has not yet been determined at the stage where \tilde{p} is found. An explicit treatment of the viscous linear terms alleviates the problem, and can be chosen to have high accuracy, but has been found to be a source of instability for the incompressible flow equations. Karniadakis [6] avoids the instability, and treats the viscous terms using the vector identity

$$\nabla^2 \mathbf{u} = \nabla (\nabla \cdot \mathbf{u}) - \nabla \times \nabla \times \mathbf{u}. \quad (63)$$

Using this idea with equation (45), the result is

$$\begin{aligned} \frac{\partial \tilde{p}^{n+1}}{\partial z} = & w^n - w^{n+1} + \sum_{q=0} \alpha_q W_w^{n+1-q} \\ & + \alpha_0 \frac{\partial}{\partial z} \nabla \cdot \mathbf{u}^{n+1} + \frac{\partial}{\partial z} \sum_{q=1} \alpha_q \nabla \cdot \mathbf{u}^{n+1-q} - \mathbf{k} \cdot \sum_{q=0} \beta_q \nabla \times \nabla \times \mathbf{u}^{n-q} - \sum_{q=0} \beta_q N_w^{n-q}, \end{aligned} \quad (64)$$

where \mathbf{k} is the unit vector. The divergence term at the leading time step can then be modified using the continuity equation,

$$\nabla \cdot \mathbf{u}^{n+1} + \frac{\bar{\rho}_z}{\bar{\rho}} w^{n+1} = 0. \quad (65)$$

by substituting for $\nabla \cdot \mathbf{u}^{n+1}$. For the incompressible flow case, this technique eliminates w^{n+1} from the pressure boundary condition. But for the anelastic equations, such a simplification does not occur, since w^{n+1} remains. The value of w^{n+1} is generally known at the boundary from the velocity boundary condition, but its derivative is not.

The anelastic equations have the wave terms, which are treated implicitly, and provide a measure of numerical stability. Due to the previous problem of the pressure boundary conditions, the viscous term in the pressure boundary condition must be treated explicitly. However this does not cause instability, apparently due to the influence of the wave terms.

The final boundary condition is

$$\frac{\partial \bar{p}^{n+1}}{\partial z} = u^n - u^{n-1} + \sum_{q=0} \alpha_q W_w^{n+1-q} - \sum_{q=0} \beta_q V_w^{n-q} + \sum_{q=0} \beta_q L_w^{n-q}. \quad (66)$$

8 Results

An example case is considered. The mean temperature profile consists of two layers, each with a different value of the Brunt-Vaisala frequency. The tropopause is located at z of 60 km.

One incident wave is considered. The horizontal wavelength is 10 km and the vertical wavelength initially is 1 km. Figures 1 and 2 show the vertical velocity versus elevation at a variety of times during the evolution of the wave. Note that the wave amplitude is small for the example in figures 1 and 2, deliberately chosen to avoid wave breaking.

Several interesting features are evident. Figure 2 shows that at a time of 2.46 hr, the wave is saturated at an elevation in the vicinity of 100 km, which indicates that a secondary flow, meaning convection or turbulence of some type is in its initial phase, and is extracting energy from the wave. Also evident is a sudden change in the vertical wavelength at an elevation of 60 km.

Figure 3 shows the vertical wavelength of the waves shown in figure 2 for a time of 2.46 hr. The wavelength was determined by finding the elevations where the vertical velocity is zero, and then a 'wave' is associated with every two of the zeros. Figure 3 clearly shows a sudden shift in the wavelength at 60 km, which is exactly the location of the tropopause. Above the tropopause, where the Brunt-Vaisala is constant, the wavelength does not change significantly. This is as expected.

Below the tropopause, the wavelength is seen to change slowly, in spite of the fact that

the Brunt-Vaisala frequency is constant here also. It is not yet clear why the wavelength changes, but it is possible that incident waves are partially reflected at the tropopause and interact nonlinearly with later incident waves, resulting in the drift of wavelength. The reflection of wave of the tropopause will be considered in future work.

References

- [1] F. B. Lipps and R. S. Hemler A scale analysis of deep moist convection and some related numerical calculations. *Journal of Atmospheric Science*, **39**, pp. 2192-2210, 1959.
- [2] R. Wilhelmson and Y. Ogura On the pressure perturbation and the numerical modeling of a cloud. *Journal of Atmospheric Science*, **29**, pp. 1295-1307, 1959.
- [3] Y. Ogura and N. A. Phillips Scale analysis of deep and shallow convection in the atmosphere. *Journal of Atmospheric Science*, **19**, pp. 173-179, 1962.
- [4] J. T. Bacmeister and M. R. Schoeberl Breakdown of vertically propagating two-dimensional gravity waves forced by orography. *Journal of Atmospheric Science*, **46**, pp. 2109-2134, 1989.
- [5] R. L. Walterscheid and G. Schubert Nonlinear evolution of an upward propagating gravity wave: overturning, convection, transience, and turbulence. *Journal of Atmospheric Science*, **47**, pp. 101-125, 1990.
- [6] G. E. Karniadakis, M. Israeli, and S. A. Orszag High-order splitting methods for the incompressible Navier-Stokes equations. *Journal of Computational Physics*, **97**, pp. 414-443, 1991.
- [7] J. R. Holton *An Introduction to Dynamics Meteorology*. Academic Press, 1992.
- [8] O. Andreassen, C. E. Wasberg, D. C. Fritts, and J. R. Isler, Gravity wave breaking in two and three dimensions 1. Model description and comparison of two-dimensional evolutions. *Journal of Geophysical Research*, **99**, pp. 8095-8108, 1994.

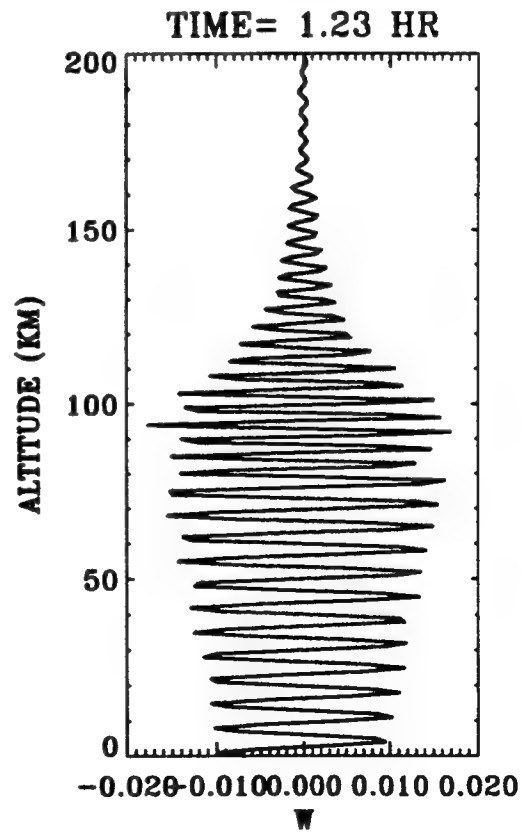
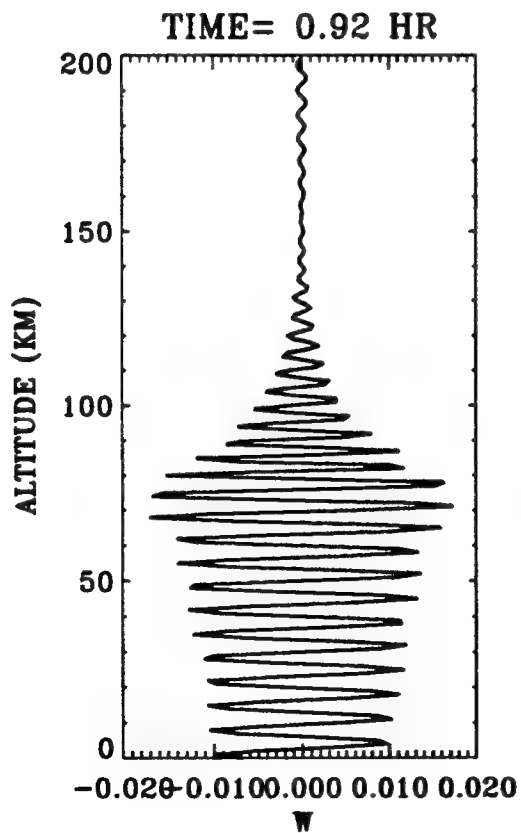
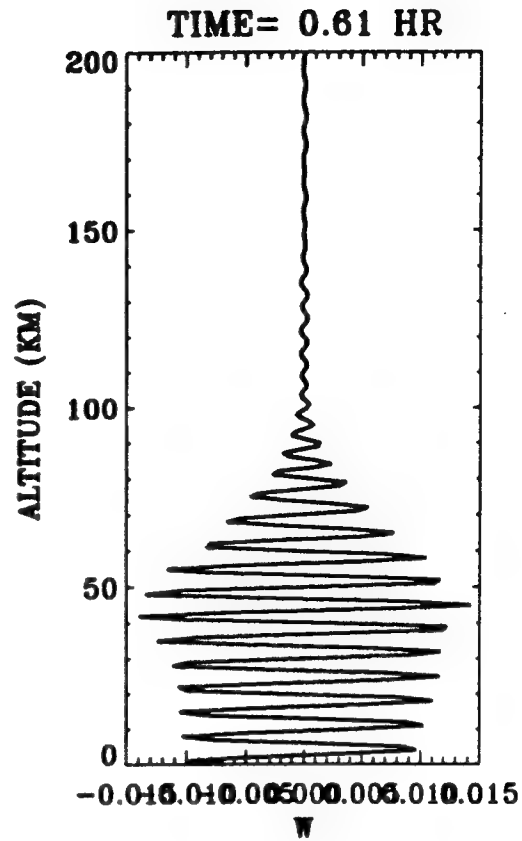
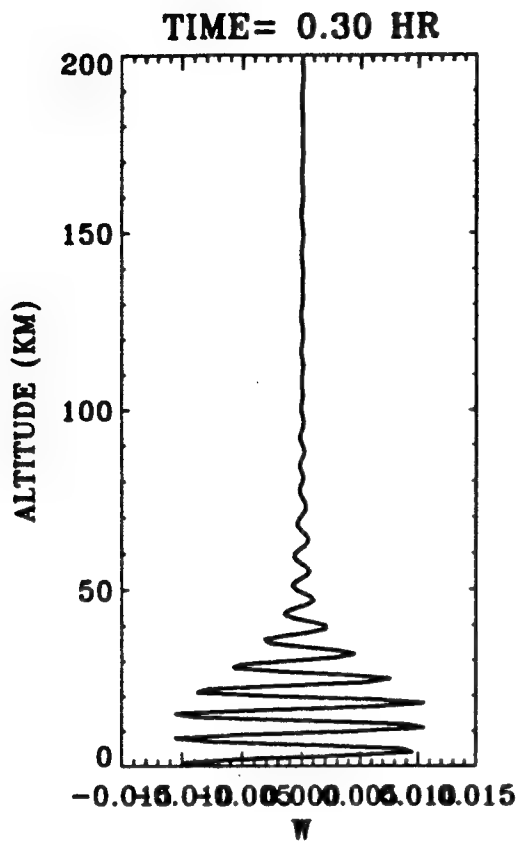


Figure 1

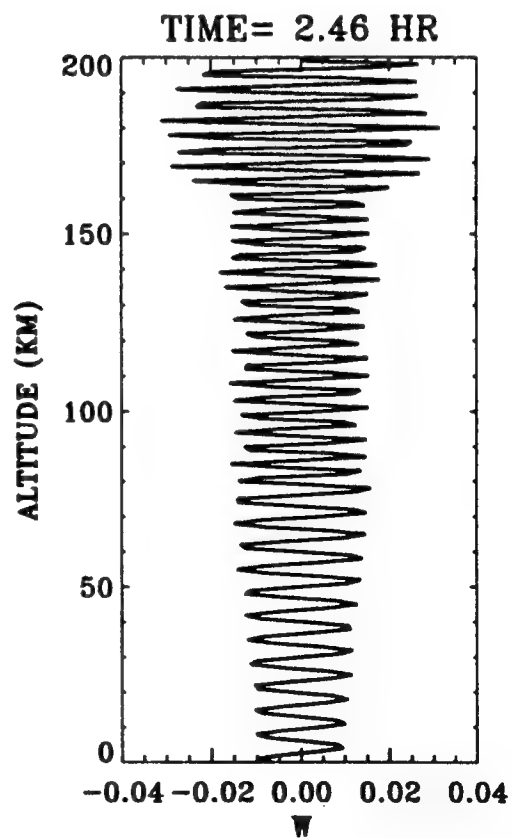
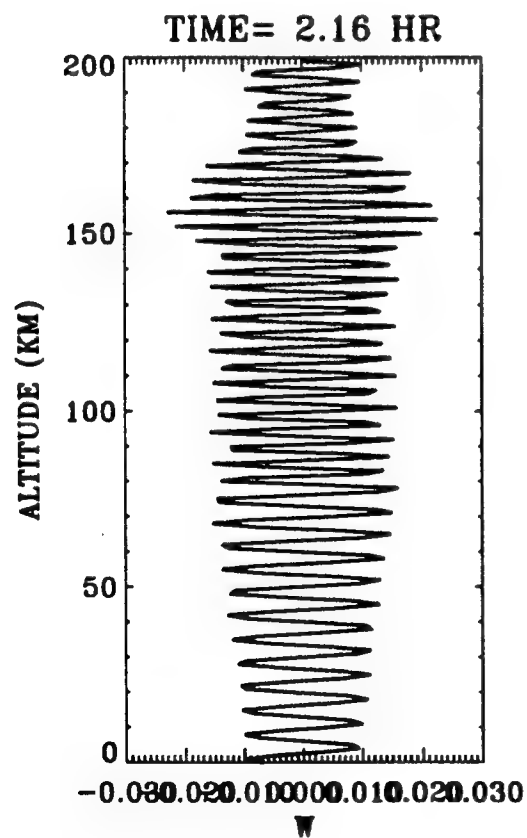
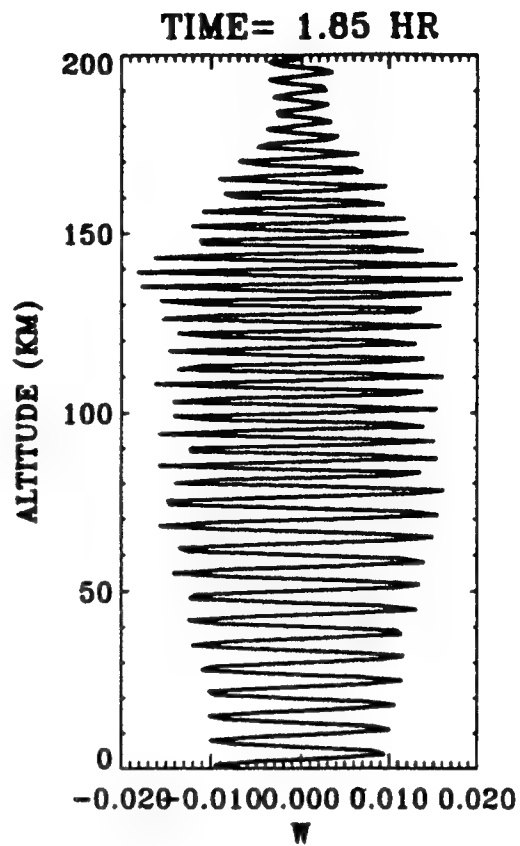
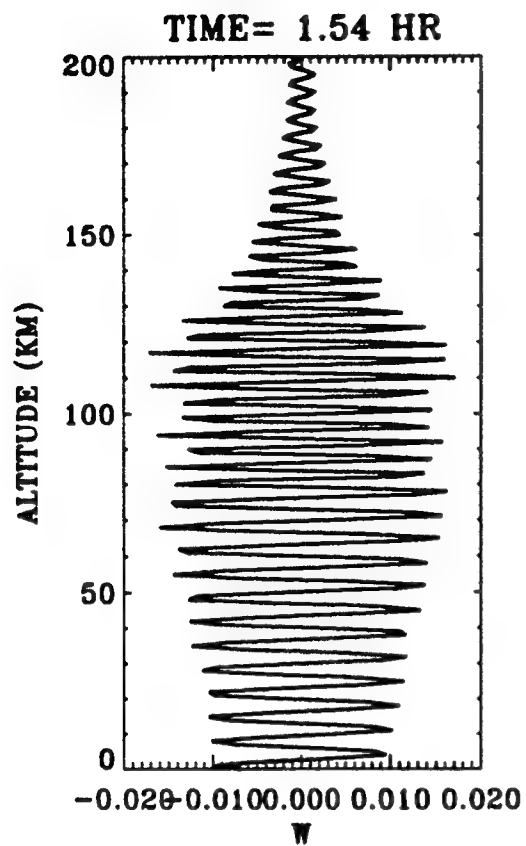


Figure 2

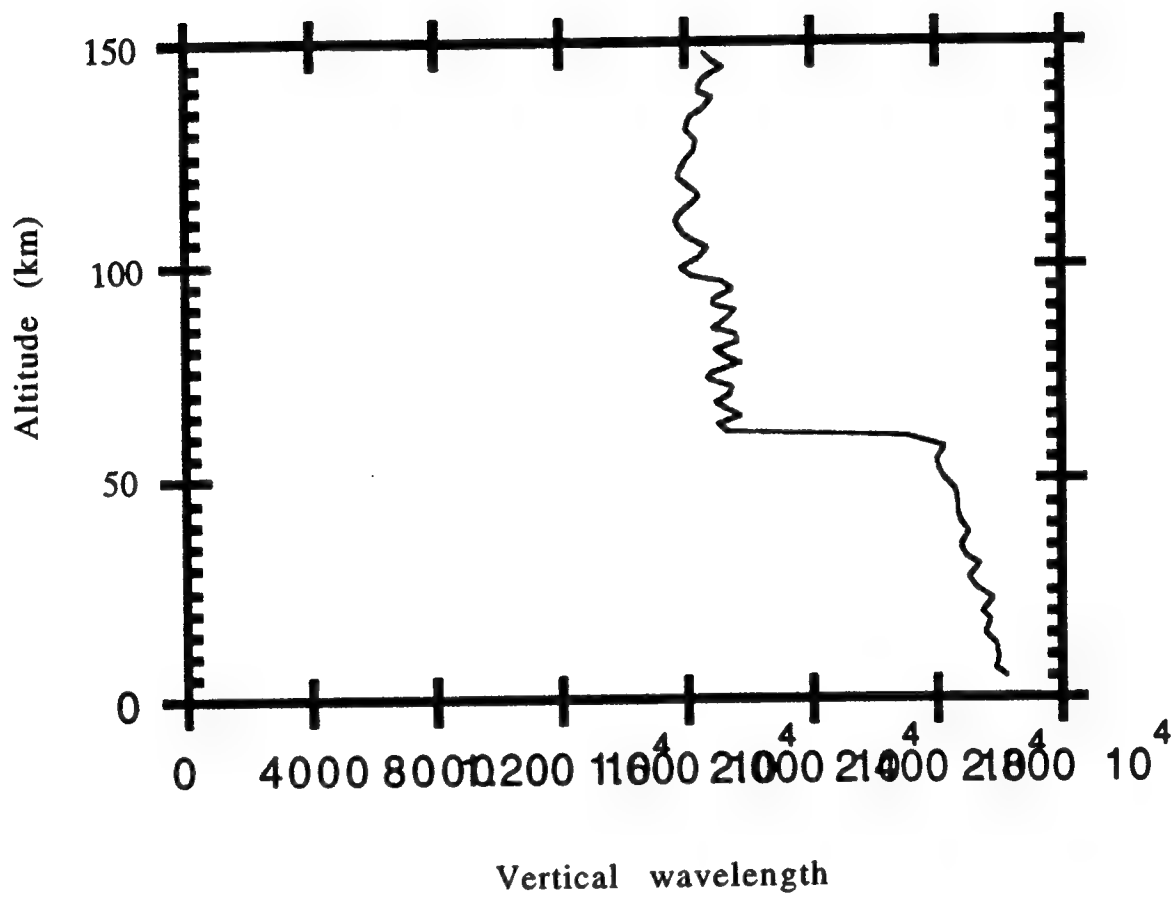


Figure 3

**LIE-ALGEBRAIC REPRESENTATIONS OF
PRODUCT INTEGRALS OF VARIABLE MATRICES**

**Stanly Steinberg
Professor of Mathematics
Department of Mathematics and Statistics**

**University of New Mexico
Albuquerque NM 87131-1141 USA
stanly@math.unm.edu
<http://math.unm.edu/~stanly/>**

**Final Report for:
Summer Faculty Research Program
Air Force Research Laboratory**

**Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC**

and

Air Force Research Laboratory

September 1998

**LIE-ALGEBRAIC REPRESENTATIONS OF
PRODUCT INTEGRALS OF VARIABLE MATRICES**

**Stanly Steinberg
Professor of Mathematics
Department of Mathematics and Statistics
University of New Mexico**

Abstract

Lie algebraic ideas are used to find useful representations of product integrals of variable matrix functions. These representations are then used to construct explicit solutions to variable coefficient differential equations that have applications to transmission-line and wave-launcher problems.

LIE-ALGEBRAIC REPRESENTATIONS OF PRODUCT INTEGRALS OF VARIABLE MATRICES

Stanly Steinberg

1 Introduction

We will use some ideas from the theory of Lie algebras and groups to find representations of solutions of systems of first order ordinary differential equations. The system will be written in the form

$$\frac{d\vec{y}}{dt} = A(t)\vec{y} + \vec{f}(t), \quad \vec{y}(0) = \vec{y}_0, \quad (1.1)$$

where $t \in [0, T]$ with $T > 0$ and $A(t)$ is a given sufficiently-smooth n by n matrix-valued function of $t \in [0, T]$, $\vec{f}(t)$ is a given sufficiently-smooth n -vector-valued function of $t \in [0, T]$, \vec{y}_0 is some given n -vector, and $n \geq 1$. It is well-known that such problems have a smooth solution $\vec{y}(t)$ for $t \in [0, T]$.

The most important representation theorem for such problems says that the solution $\vec{y}(t)$ of (1.1) can be written in the form

$$\vec{y}(t) = G(t, 0) \vec{y}_0 + \int_0^t G(t, \tau) d\tau, \quad 0 \leq t \leq T, \quad (1.2)$$

where $G(t, \tau)$ is a smooth n by n matrix-valued solution of

$$\frac{d}{dt}G(t, \tau) = A(t)G(t, \tau), \quad G(\tau, \tau) = I, \quad 0 \leq \tau \leq t \leq T, \quad (1.3)$$

where I is the n by n identity matrix. The function G is called the *Green's function*, *propagator*, or *matrizant*. These results can be found in many textbooks on ordinary-differential equations. The book by [6] by Dollard and Friedman reviews this material and additionally is the primary reference for some material we will use later.

In the constant coefficient case, that is, when $A(t) = A$ is independent of t , then

$$G(t, \tau) = e^{A(t-\tau)}, \quad (1.4)$$

where for any matrix M

$$e^M = \sum_{n=0}^{\infty} \frac{M^n}{n!}. \quad (1.5)$$

This series converges absolutely and uniformly in the entries of M , and can thus be differentiated term by term. So setting $M = At$ gives

$$e^{A0} = I, \quad \frac{d}{dt}e^{At} = Ae^{At}, \quad (1.6)$$

which implies that (1.4) is in fact the Green's function. If the matrix function $A(t)$ is not constant then Green's function G can still be represented in an exponential form using the product integral as described in Dollard and Friedman:

$$G(t, \tau) = \prod_{\tau}^t e^{ds A(s)}, \quad (1.7)$$

where the product integral can be defined as

$$\prod_{\tau}^t e^{ds A(s)} = \lim_{h \rightarrow 0} \prod_{i=0}^{N-1} e^{h A(\tau + ih)}, \quad (1.8)$$

where $h = (t - \tau)/N$, $A(t)$ is a continuous function of t , and the product in 1.8 means *left dot* multiplication by successive term (increasing i).

The main point we want to take from this introductory section is that the solution of systems of ordinary differential equations can be written in terms of exponentials of matrices, even in the variable coefficient case, where product integrals must be used. The theory and formulas presented here are quite useful, but our goals is to produce much more useful formulas using the notions of Lie algebras and groups.

This project resulted from the authors realizing that there was a close connection between some of the second authors work on transmission lines and wave launchers [2, 3, 4] and the first authors research on the applications of Lie methods to differential equations [9, 10, 11, 12, 13].

2 Lie Algebras and Lie Groups

The text [14] has a useful introduction to Lie algebra and groups, and Chapters 2 and 8 of the book [1] edited by Baum and Kritikos have nice overviews of Lie algebras and groups and their connection to Electromagnetics. However, we will include here a complete discussion of all of the facts that we need. We will restrict our attention to Lie algebras and groups of n by n matrices with entries from either the complex field or the real field. A *Lie algebra* of matrices is a non-empty set of matrices that is closed under addition, scalar multiplication and commutation. The commutator of two matrices is defined by

$$[A, B] = AB - BA. \quad (2.9)$$

A *Lie group* of matrices is a non-empty set of matrices that is closed in multiplication and matrix inversion along with an additional property that will be described later.

The set of all n by n matrices is a Lie algebra called $gl(n)$ which stands for the *general linear algebra*. The general linear algebra $gl(n)$ is a linear space with addition being the standard addition of entries in the matrix and scalar multiplication being multiplication of all entries by an element of the field. The matrices with non-zero determinant form a Lie group called the *general linear group*. These examples will be our main algebras and groups with all other examples being subsets of these.

In applications to differential equations, the matrices $A(t)$ in (1.1) will be taken from a Lie algebra, while the Green's function, propagator, or matrizant used in (1.3) will be in a Lie group.

If $A(t) = A$ is constant, then the solutions of the system of ordinary differential equations (1.1) can be written in terms of the Green's function $G(t, \tau) = G(t - \tau)$ where

$$G(t) = e^{At}. \quad (2.10)$$

The Lie group associated with a given Lie algebra is supposed to represent the set of all possible Green's functions generated by elements of the Lie algebra. More precisely, the Lie group associated with a Lie algebra is the set of all possible finite products of exponentials of elements from the Lie algebra.

To understand Lie algebras and groups better, we need a few facts about exponential of matrices. Recall that the exponential is defined by

$$e^{At} = \sum_{n=0}^{\infty} \frac{A^n t^n}{n!}, \quad (2.11)$$

which is a uniformly and absolutely convergent series for all t , and consequently the entries in the $\exp(At)$ are all infinitely differentiable. It is an easy to check that

$$\begin{aligned} e^{A0} &= I, \\ e^{At} e^{-At} &= I \\ (e^A)^\dagger &= e^{A^\dagger} \\ e^{A(t_1+t_2)} &= e^{At_1} e^{At_2} = e^{At_2} e^{At_1}, \\ \frac{d}{dt} e^{At} &= A e^{At} = e^{At} A, \end{aligned} \quad (2.12)$$

where A^\dagger is the complex conjugate transpose of the matrix A . In particular, the second identity implies that the matrix $\exp(At)$ is invertible with inverse $\exp(-At)$.

An important tool for evaluating exponentials is the formula for how the exponential changes under a similarity transformation:

$$e^{S^{-1}AS} = S^{-1}e^A S, \quad \det(S) \neq 0, \quad (2.13)$$

which is easy to check from the series definition of the exponential.

It is easy to evaluate exponentials if the matrix can be put into a simple block form. For example, if A is block matrix of the form

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} \quad (2.14)$$

then

$$A^n = \begin{pmatrix} A_1^n & 0 \\ 0 & A_2^n \end{pmatrix} \quad (2.15)$$

and then the series definition of the exponential gives

$$e^{At} = \begin{pmatrix} e^{A_1 t} & 0 \\ 0 & e^{A_2 t} \end{pmatrix}. \quad (2.16)$$

In particular, if A is a diagonal matrix, $A = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, then

$$e^{At} = \text{diag}(e^{\lambda_1 t}, e^{\lambda_2 t}, \dots, e^{\lambda_n t}). \quad (2.17)$$

For any matrix there is a similarity transformation that reduces the matrix to Jordan form, that is, the matrix is block diagonal with Jordan blocks on the diagonal. So to compute the exponential of any matrix, we only need to know how to compute the exponentials of the Jordan blocks. Unfortunately, in general, it is difficult to find the Jordan form if a matrix contains symbolic parameters or floating-point numbers. Still, this is a useful tool for small matrices. The simplest Jordan block is 2 by 2:

$$J = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \quad (2.18)$$

and, using the power series definition, its exponential is easily computed to be

$$e^{Jt} = e^{\lambda t} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}. \quad (2.19)$$

Another way to find this result is to convert the computation of the exponential to a system of differential equations. So assume that

$$e^{Jt} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad (2.20)$$

where $a = a(t)$, $b = b(t)$, $c = c(t)$, $d = d(t)$. Differentiating gives

$$\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \quad (2.21)$$

So a , b , c , and d satisfy the differential equations

$$a' = \lambda a + c, \quad b' = \lambda b + d, \quad c' = \lambda c, \quad d' = \lambda d, \quad (2.22)$$

and the fact that the exponential is the identity at $t = 0$ gives the initial conditions

$$a(0) = 1, \quad b(0) = 0, \quad c(0) = 0, \quad d(0) = 1. \quad (2.23)$$

The solutions are easily seen to be

$$a(t) = e^{\lambda t}, \quad b(t) = t e^{\lambda t}, \quad c(t) = 0, \quad d(t) = e^{\lambda t}, \quad (2.24)$$

which agrees with the power-series computation.

Let us now return to the problem of identifying the Lie group generated by exponentials of elements from the general linear algebra $gl(n)$. It is well known [7] that the trace $\text{tr}(A)$ of a matrix A and the determinant $\det(A)$ are related by:

Lemma:

$$\det(e^A) = e^{\text{tr}(A)}. \quad (2.25)$$

Proof: Even though well-known, this is so important for us that we give a proof. It is also well known [7] that the determinant and trace of a matrix are invariant under similarity transforms, so if $\det(S) \neq 0$, then

$$\det(S^{-1} A S) = \det(A), \quad \text{tr}(S^{-1} A S) = \text{tr}(A). \quad (2.26)$$

For matrices in Jordan form (diagonal form, or even triangular form) it is simple to see that

$$\det(e^J) = e^{\text{tr}(J)}. \quad (2.27)$$

Any matrix can be put into Jordan form, so for any A there is an S so that $J = S^{-1} A S$ and then

$$\det(e^A) = \det(S^{-1} e^A S) = \det(e^{S^{-1} A S}) = \det(e^J) = e^{\text{tr}(J)} = e^{\text{tr}(S J S^{-1})} = e^{\text{tr}(A)}. \quad (2.28)$$

When the matrix has complex entries, the trace is an exponential of a complex number which cannot be zero, so the matrix is invertible. In fact, we already know $\exp(A t) = \exp(-A t)$ (see (2.12)). When the entries of a matrix are real, then the trace is real, and because the exponential function is always positive, the exponential of a matrix always has a positive determinant and thus is invertible. So the group $gl(n)$ is the set of matrices with positive determinant.

We digress for a moment to consider the more difficult question: can every matrix M with positive determinant be written as an exponential? Clearly we should take the logarithm of M . If $|x| < 1$, then the base e logarithm of $1 + x$ can be expressed as a power series:

$$\ln(1 + x) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^n}{n}. \quad (2.29)$$

The size of a matrix can be specified by giving a norm on all matrices. We will use the norm that is the square root of the sum of the squares of the entries of the matrix, called the *root span* which can be conveniently written as

$$\|A\| = \text{rsp}(A) = \sqrt{\text{tr}(A A^\dagger)}, \quad (2.30)$$

where A^\dagger is the complex conjugate transpose of the matrix A . So if $\|M\| < 1$, then we define

$$\ln(I + M) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{(M)^n}{n}. \quad (2.31)$$

An elementary power series calculation shows that for x sufficiently small x

$$e^{\ln(1+x)} = 1 + x, \quad \ln(e^x) = x, \quad (2.32)$$

so that for M sufficiently small

$$e^{\ln(1+M)} = 1 + M, \quad \ln(e^M) = M, \quad (2.33)$$

Because any matrix near the identity matrix can be written as $I + M$ where M is small, we conclude that all matrices with non-zero determinant that are sufficiently near the identity can be written as the exponential

of another matrix, in general with complex entries. This is definitely not true for all matrices with non-zero determinant. Because the logarithm has a limited radius of convergence and is generally hard to work with, the use of the logarithm to define the Lie algebra given a Lie group is not such a good idea.

To study the Lie algebra associated with a Lie group, it is better to use differentiation, in particular, the logarithmic derivative. Let $G(t)$ be a function from some interval to the group and then define the elements of the Lie algebra to be the *product derivative*

$$A = D_t G(t)|_{t=a} = \left\{ \frac{d}{dt} G(t) \right\} G^{-1}(t) \Big|_{t=a}. \quad (2.34)$$

for all a in the interval and all possible functions $G(t)$. Actually one can restrict the interval to contain zero, the functions to have $G(0) = I$ and then

$$A = \frac{d}{dt} G(t) \Big|_{t=0} \quad (2.35)$$

are sufficient to give all elements in the algebra. Now this is nice as we see that if A is in a Lie algebra, then $G(t) = \exp(A t)$ is in the Lie group and then

$$\frac{d}{dt} G(t) \Big|_{t=0} = \frac{d}{dt} e^{A t} \Big|_{t=0} = A e^{A t} \Big|_{t=0} = A, \quad (2.36)$$

or the logarithmic derivative gives

$$\frac{d}{dt} G(t) G^{-1}(t) = \frac{d}{dt} e^{A t} e^{-A t} = A e^{A t} e^{-A t} = A, \quad (2.37)$$

so things are nicely consistent. The additional assumption in the definition of a Lie group mentioned earlier can be stated as: if G belongs to the Lie group, then there is a function $G(t)$ with $G(t)$ in the group and $G(0) = I$ and $G(1) = G$. This actually defines a connected Lie Group, for example, for matrices with real entries, the natural Lie group consists of the matrices with positive determinant (not just non-zero determinant).

So given a lie group the Lie algebra can be found by computing derivatives at the origin. Given a Lie algebra, the Lie group can be found by computing exponentials of the elements of the Lie algebra.

Now all of this will give us some more examples of Lie algebras and groups which we have summarized in Table 2.1. The general linear algebras and groups have already been discussed. Because the trace is a linear operation it is clear that the traceless matrices are preserved under scalar multiplication and addition. If $A = (a_{i,j})$ and $B = (b_{i,j})$ then

$$\text{tr}(A B) = \sum_{i,j=1}^n a_{i,j} b_{j,i} = \text{tr}(B A), \quad (2.38)$$

so for all matrices

$$\text{tr}([A, B]) = \text{tr}(A B - B A) = \text{tr}(A B) - \text{tr}(B A) = 0. \quad (2.39)$$

In particular, the set of traceless matrices is closed under both linear operations and commutation, and so they form a lie algebra. Because the exponential of a traceless matrix must have determinant one due to formula (2.25), the Lie group associated with the traceless matrices are the matrices of determinant one.

Next we will show that the group of unitary matrices, that is, matrices that satisfy

$$M^\dagger M = I \quad (2.40)$$

have the associated Lie algebra of skew symmetric matrices, that is, matrices that satisfy

$$A^\dagger = -A. \quad (2.41)$$

algebra		group	
$gl(n)$	all A	$GL(n)$	$\det(M) \neq 0$
$sl(n)$	$\text{tr} A = 0$	$SL(n)$	$\det(M) = 1$
$u(n)$	$A^\dagger = -A$	$U(n)$	$M^\dagger M = I$

Table 2.1: Lie Algebras and Group

To show that the group associated with the algebra is correct, assume that $M = \exp(At)$ with $A^\dagger = -A$. Then (2.12) gives

$$M^\dagger M = e^{A^\dagger t} e^{At} = e^{-At} e^{At} = I, \quad (2.42)$$

so the exponential of a skew adjoint matrix is unitary. To show that the algebra associated with the group is correct, suppose that $M(t)$ is a matrix function satisfying

$$M^\dagger(t) M(t) = I, \quad (2.43)$$

so

$$M^\dagger(t) = M^{-1}(t), \quad (2.44)$$

and then also

$$M(t) M^\dagger(t) = I, \quad (2.45)$$

Now set

$$A = M'(t) M^{-1}(t) = M'(t) M^\dagger(t) \quad (2.46)$$

and then differentiate (2.43) to get

$$M'^\dagger(t) M(t) + M^\dagger(t) M'(t) = 0, \quad (2.47)$$

Multiplying on the left by M and on the right by M^\dagger gives

$$M(t) M'^\dagger(t) + M'(t) M^\dagger(t) = 0, \quad (2.48)$$

or

$$A^\dagger + A = 0. \quad (2.49)$$

So, A is skew adjoint, and the Lie algebra associated with the unitary matrices consists of the skew symmetric matrices and vice versa. Similar arguments can be used to find the connection between various Lie groups and Lie algebras. These results are summarized in Table 2.1. For more examples, see [14] or any text on Lie algebras and groups.

3 Lie Theory and Ordinary Differential Equations

In this section, we show how to use Lie theory to solve systems of ordinary differential equations. We begin with the derivation of two important formulas and then describe the "solution" algorithm.

3.1 Important Lie Formulas

To apply Lie theory to differential equations, we need a few special formulas. The first is
Lemma:

$$\frac{d}{dt} e^{f(t)A} = f'(t) A e^{f(t)A}, \quad (3.50)$$

where $f(t)$ is any smooth function and A is any constant square matrix.

This is easily checked using the power series definition (2.11) of the exponential.

The second formula is for the *adjoint action* of the Lie group on the Lie algebra, which appears repeatedly when formulas are simplified or rearranged. First, for any two square matrices A and B of the same size, define

$$[A, \circ]^0 B = B, \quad [A, \circ]^n = [A, [A, \circ]^{n-1}], \quad (3.51)$$

so that

$$\begin{aligned} [A, \circ]^1 B &= [A, B], \\ [A, \circ]^2 B &= [A, [A, B]], \\ [A, \circ]^3 B &= [A, [A, [A, B]]], \end{aligned}$$

and so forth. Now the adjoint action of A on B is given by

$$e^{[A, \circ]} B = \sum_{n=0}^{\infty} \frac{[A, \circ]^n B}{n!}. \quad (3.52)$$

Lemma:

$$e^A B e^{-A} = e^{[A, \circ]} B. \quad (3.53)$$

Proof: This is really a formula for rearranging power series, and can be seen to be true by comparing the power series of both sides of the equation. However, an ordinary differential equation proof is more illuminating and points out the power of this "trick". Let

$$F(t) = e^{At} B e^{-At} - e^{t[A, \circ]} B. \quad (3.54)$$

Then

$$\begin{aligned} F'(t) &= A e^{At} B e^{-At} - e^{At} B e^{-At} A - [A, \circ] e^{t[A, \circ]} B \\ &= [A, e^{At} B e^{-At}] - [A, e^{t[A, \circ]} B] \\ &= [A, F(t)]. \end{aligned} \quad (3.55)$$

So

$$F'(t) = [A, F(t)], \quad F(0) = 0, \quad (3.56)$$

which implies that $F(t) \equiv 0$.

Note that if A and B commute $[A, B] = 0$, then

$$e^{At} B e^{-At} = e^{t[A, \circ]} B = B. \quad (3.57)$$

3.2 Solution Algorithm

Recall that the system of ordinary differential equations (1.1) is determined by the matrix function $A(t)$. Because the $A(t)$ are n by n matrices, they certainly belong to the Lie algebra $GL(n)$. The first step in the solution process is to find the smallest Lie algebra \mathcal{A} of matrices into which all of the matrices $A(t)$ fall.

The second step is to choose a basis for the Lie algebra \mathcal{A} , say

$$\mathcal{A} = \text{span}(A_1, A_2, \dots, A_k), \quad 1 \leq k \leq n^2, \quad (3.58)$$

where here *span* means all real or complex linear combinations of the matrix bases elements A_i , that is

$$A = \sum_{i=1}^k a_i(t) A_i, \quad (3.59)$$

where $a_i(t)$ are either real or complex valued functions of t .

Now the previous discussion implies that for any real numbers a_i , $1 \leq i \leq k$, the matrices

$$M(a_1, a_2, \dots, a_k) = e^{a_1 A_1 + a_2 A_2 + \dots + a_k A_k} . \quad (3.60)$$

are in the Lie group and all of the matrices near the identity in the Lie group can be written this way. The numbers a_i are called the coordinates on the group near the identity. This is the typical mathematical approach to generating Lie groups from Lie algebras. However there are better ways to represent a Lie group if we are interested in finding simple formulas for the representation, which is:

$$M(a_1, a_2, \dots, a_k) = e^{a_1 A_1} e^{a_2 A_2} \dots e^{a_k A_k} . \quad (3.61)$$

So we have exchanged an exponential of a sum for a product of exponentials. Again it is clear that G is in the Lie group, but is also true that all elements of the group near the identity can be represented in this way. There are many possible variant of this representation. Also note that Lie algebras have many bases and the basis impacts the simplicity of the resulting formulas. One good idea is to choose as many elements of the basis as possible that commute with each other, that is, choose part of the basis as a basis for the largest commuting sub-algebra.

For any computation, a table of commutators

$$[A_i, A_j], \quad 1 \leq i, j \leq k \quad (3.62)$$

will be needed, so the third step is to compute this table.

Recall that the solution of the differential equations is given in terms of G which is the Green's function, propagator, or matrizant, which is the solution of the system of differential equations (1.3):

$$\frac{d}{dt} G(t, \tau) = A(t) G(t, \tau), \quad G(\tau, \tau) = I, \quad 0 \leq \tau \leq t \leq T. \quad (3.63)$$

So Lie theory tells us that the Green's function can be written in the form

$$G(t, \tau) = e^{g_1(t, \tau) A_1} e^{g_2(t, \tau) A_2} \dots e^{g_k(t, \tau) A_k} . \quad (3.64)$$

We call this the *fundamental* representation. Note that it depends strongly on the basis and on the ordering of the basis.

First note that because we must have $G(\tau, \tau) = I$, we need

$$g_i(\tau, \tau) = 0, \quad 1 \leq i \leq k. \quad (3.65)$$

Next we want to plug the representation (3.64) into the differential equation in (1.3), so we need to differentiate the representation. Applying the product rule and the derivative of exponentials formula (3.50) gives

$$\begin{aligned} \frac{d}{dt} G(t, \tau) &= g'_1(t, \tau) A_1 e^{g_1(t, \tau) A_1} e^{g_2(t, \tau) A_2} e^{g_3(t, \tau) A_3} \dots \\ &\quad + e^{g_1(t, \tau) A_1} g'_2(t, \tau) A_2 e^{g_2(t, \tau) A_2} e^{g_3(t, \tau) A_3} \dots \\ &\quad + e^{g_1(t, \tau) A_1} e^{g_2(t, \tau) A_2} g'_3(t, \tau) A_3 e^{g_3(t, \tau) A_3} \dots \\ &\quad + \dots \end{aligned} \quad (3.66)$$

Now we can plug in the derivative of G into the differential equation and then multiply the resulting equation on the *right* by the inverse of G to get the logarithmic, product, or multiplicative derivatives (see Dollard and Friedman [6], Section 1.3) on the left-hand side of the equation. The inverse of G can be written in the form

$$G^{-1}(t, \tau) = e^{-g_k(t, \tau) A_k} \dots e^{-g_2(t, \tau) A_2} e^{-g_1(t, \tau) A_1} . \quad (3.67)$$

We note that the left-hand side of the resulting equation is a logarithmic derivative:

$$\frac{d}{dt} G(t, \tau) G^{-1}(t, \tau) = A(t), \quad (3.68)$$

while the right-hand side has many exponentials that cancel giving the *determining* equation

$$\begin{aligned} A(t) = & g'_1(t, \tau) A_1 \\ & + g'_2(t, \tau) e^{g_1(t, \tau) A_1} A_2 e^{-g_1(t, \tau) A_1} \\ & + g'_3(t, \tau) e^{g_1(t, \tau) A_1} e^{g_2(t, \tau) A_2} A_3 e^{-g_2(t, \tau) A_2} e^{-g_1(t, \tau) A_1} \\ & + \dots \end{aligned} \quad (3.69)$$

This equation along with the initial conditions (3.65) determine the g_i functions.

It is always illuminating to look at any formula when all of the basis elements commute, $[A_i, A_j] = 0$, $1 \leq i, j \leq k$. In this case the determining equations are

$$A(t) = a_1(t) A_1 + a_2(t) A_2 + \dots + a_k(t) A_k = g'_1 A_1 + g'_2 A_2 + \dots + g'_k A_k. \quad (3.70)$$

So

$$g'_i(t, \tau) = a_i(t), \quad g_i(\tau, \tau) = 0, \quad 1 \leq i \leq k, \quad (3.71)$$

or

$$g_i(t, \tau) = \int_{\tau}^t a_i(s) ds, \quad 1 \leq i \leq k. \quad (3.72)$$

In particular, if the a_i are constant, $a_i(t) = a_i$, then

$$g_i(t, \tau) = a_i(t - \tau), \quad 1 \leq i \leq k. \quad (3.73)$$

Finally, the matrizant has the form (3.64) where the order of the factors doesn't matter. Baum has already made significant use of this special representation, see [4], Appendix C.

The right-hand side of the determining equation can be evaluated using the adjoint-action formula (3.53), but to do this we need a table of the adjoint actions:

$$F_{i,j}(t) = e^{A_i t} A_j e^{-A_i t}, \quad 1 \leq i, j \leq k, \quad (3.74)$$

These adjoint actions can be evaluated using either the series in (3.53) or by solving the initial-value problems

$$\frac{d}{dt} F_{i,j}(t) = [A_i, F_{i,j}(t)], \quad F_{i,j}(0) = A_j. \quad (3.75)$$

We now have enough theory to construct the functions g_i in the fundamental representation (3.64) of the Green's function. To finish, we need a table of the exponentials of the basis elements

$$G_i(t) = e^{A_i t}. \quad (3.76)$$

Again these can be evaluated by either using the series definition of the exponential or the fact that the G_i satisfy the initial-value problems

$$\frac{d}{dt} G_i(t) = A_i G_i(t), \quad G_i(0) = I. \quad (3.77)$$

The resulting form of the determining equation is the equality of two expressions which are linear combinations of the basis elements which says that the coefficient of the basis elements must be equal, which yields k (the number of basis elements) equations in k unknowns g_i . These equations can be solved for g'_i which yields k nonlinear differential equations. Now this seems like a bad deal. We started with n^2 linear variable-coefficient differential equations and end up with $k \leq n^2$ non-linear constant coefficient differential equations (which are difficult to solve even in the simplest cases). However, we have done this in such a way that if we choose the g_i as arbitrary functions, then we produce an $A(t)$ which gives a system that is exactly solvable with the solution of the homogeneous equations given by a known Green's function $G(t, 0)$. As we will see, this can be used to great advantage in studying physical systems.

4 Examples

In this section we will look at some one, two and three dimensional examples. Note that many of the computations were done using a computer algebra system. None of the computations are difficult, just that there are quite a few elementary computations. The computer algebra programs are available from S. Steinberg.

4.1 One-Dimensional Example

One dimensional examples are scalar equations whose Green's function satisfies

$$\frac{d}{dt}G(t, \tau) = A(t)G(t, \tau), \quad G(\tau, \tau) = 1. \quad (4.78)$$

The solution of this initial value problem is

$$G(t, \tau) = e^{\int_{\tau}^t a(s) ds}. \quad (4.79)$$

One view of our job is to find formulas like this for higher-dimensional examples.

4.2 Two-Dimensional Examples

We now do a complete analysis of systems of two ordinary differential equations, that is, we assume that

$$A(t) = \begin{pmatrix} a(t) & b(t) \\ c(t) & d(t) \end{pmatrix}, \quad (4.80)$$

and that $A(t)$ is a real matrix, but similar results are true for complex matrices.

4.2.1 The Obvious Basis

The first step is to choose a basis for the 2 by 2 matrices. An obvious simple basis is

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \quad (4.81)$$

The matrix A is represented in this basis as

$$A(t) = a(t)A_1 + d(t)A_2 + b(t)A_3 + c(t)A_4. \quad (4.82)$$

The formula for the Green's function involves the exponential of the basis elements, so we will need these:

$$e^{tA_1} = \begin{pmatrix} e^t & 0 \\ 0 & 1 \end{pmatrix}, \quad e^{tA_2} = \begin{pmatrix} 1 & 0 \\ 0 & e^t \end{pmatrix}, \quad e^{tA_3} = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}, \quad e^{tA_4} = \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix}. \quad (4.83)$$

These formulas can be easily evaluated in two different ways: one is to compute a few terms of the power series definition and then identify the resulting series; the other is to convert the evaluation to an initial-value problem. For example, the power-series method gives

$$e^{tA_1} = \begin{pmatrix} 1 + t + t^2/2 + t^3/6 + t^4/24 + t^5/120 + \cdots & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} e^t & 0 \\ 0 & 1 \end{pmatrix}, \quad (4.84)$$

while if we set

$$X(t) = e^{tA_1} = \begin{pmatrix} r(t) & s(t) \\ u(t) & v(t) \end{pmatrix}, \quad (4.85)$$

then

$$X'(t) = A_1 X(t), \quad X(0) = I. \quad (4.86)$$

or

$$\begin{pmatrix} r'(t) & s'(t) \\ u'(t) & v'(t) \end{pmatrix} = \begin{pmatrix} r(t) & s(t) \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} r(0) & s(0) \\ u(0) & v(0) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (4.87)$$

So $r(t) = e^t$, $s(t) = 0$, $u(t) = 0$ and $v(t) = 1$, as was found from the power series method.

The commutator table for the basis, with entries $[A_i, A_j]$, is given by

$A_i \quad A_j$	A_1	A_2	A_3	A_4
A_1	0	0	A_3	$-A_4$
A_2	0	0	$-A_3$	A_4
A_3	$-A_3$	A_3	0	$A_1 - A_2$
A_4	A_4	$-A_4$	$-A_1 + A_2$	0

(4.88)

which is fairly simple. Next we need a table of the adjoint actions $e^{A_i t} A_j e^{-A_i t}$ of the Lie group on the Lie algebra:

$A_i \quad A_j$	A_1	A_2	A_3	A_4
A_1	A_1	A_2	$e^t A_3$	$e^{-t} A_4$
A_2	A_1	A_2	$e^{-t} A_3$	$e^t A_4$
A_3	$A_1 - t A_3$	$A_2 + t A_3$	A_3	$t A_1 - t A_2 - t^2 A_3 + A_4$
A_4	$A_1 + t A_4$	$A_2 - t A_4$	$-t A_1 + t A_2 + A_3 - t^2 A_4$	A_4

(4.89)

This was computed using the series definition of the adjoint action and then identifying the series in the resulting formulas but can also be done by converting the evaluation to a initial-value problem for a system of ordinary differential equations.

Now we represent the Green's function of the the differential equation in the form (3.64):

$$G(t, \tau) = e^{g_1(t, \tau) A_1} e^{g_2(t, \tau) A_2} e^{g_3(t, \tau) A_3} e^{g_4(t, \tau) A_4}, \quad (4.90)$$

where the g_i satisfy the initial conditions (3.65). Then the formulas for the exponentials of the basis elements (4.83) give the Green's function as a product of four matrices:

$$G(t, \tau) = \begin{pmatrix} e^{g_1} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & e^{g_2} \end{pmatrix} \begin{pmatrix} 1 & g_3 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ g_4 & 1 \end{pmatrix}. \quad (4.91)$$

The determining equation (3.69), which are computed by taking the logarithmic derivative of the Green's function and then using the table of adjoint actions, is

$$A(t) = (g'_1 + g_3 g'_4) A_1 + (g'_2 - g_3 g'_4) A_2 + (e^{g_1 - g_2} g'_3 - e^{g_1 - g_2} g_3^2 g'_4) A_3 + e^{-g_1 + g_2} g'_4 A_4. \quad (4.92)$$

Next we solve equations (4.92) with $A(t)$ given by (4.80) for g'_i to obtain the following nonlinear system of differential equations for the g_i :

$$\begin{aligned} g'_1 &= a - c e^{g_1 - g_2} g_3, \\ g'_2 &= d + c e^{g_1 - g_2} g_3, \\ g'_3 &= b e^{-g_1 + g_2} + c e^{g_1 - g_2} g_3^2, \\ g'_4 &= c e^{g_1 - g_2}. \end{aligned} \quad (4.93)$$

Even when $A(t)$ is constant, that is, a , b , c , and d are constants, the previous ordinary differential equations are seriously nonlinear and difficult to solve. However, we can create many solvable examples by

choosing g_i as some function of t and τ . If we only consider the initial value problem without source terms and choose

$$g_1(t) = \alpha t, \quad g_2(t) = \beta t, \quad g_3(t) = \gamma t, \quad g_4(t) = \delta t, \quad (4.94)$$

then the determining equations (4.92) become

$$A(t) = \begin{pmatrix} \alpha + \delta \gamma t & e^{(\alpha-\beta)t} \gamma - \delta e^{(\alpha-\beta)t} \gamma^2 t^2 \\ \delta e^{(-\alpha+\beta)t} & \beta - \delta \gamma t \end{pmatrix}. \quad (4.95)$$

If we assume the ordinary differential equation doesn't have a forcing term, then we only need $G(t) = G(t, 0)$ which is given by (4.91):

$$G(t) = \begin{pmatrix} e^{\alpha t} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & e^{\beta t} \end{pmatrix} \begin{pmatrix} 1 & \gamma t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \delta t & 1 \end{pmatrix} = \begin{pmatrix} e^{\alpha t}(1 + \delta \gamma t^2) & e^{\alpha t} \gamma t \\ \delta e^{\beta t} t & e^{\beta t} \end{pmatrix}, \quad (4.96)$$

In the applications that interest us, it is common for $A(t)$ to have trace zero or be skew symmetric (which implies trace zero). As none of our basis elements are skew symmetric and only two elements, not three, have trace zero, these situations are hard to analyze. The next basis we choose will correct this.

4.2.2 Second Basis

We now choose a basis with three trace zero matrices and one skew-symmetric matrix:

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \quad (4.97)$$

The matrix A given in (4.80) is represented in this basis as

$$A(t) = \frac{a(t) + d(t)}{2} A_1 + \frac{a(t) - d(t)}{2} A_2 + \frac{b(t) + c(t)}{2} A_3 + \frac{b(t) - c(t)}{2} A_4. \quad (4.98)$$

The exponentials of the basis elements are

$$e^{tA_1} = \begin{pmatrix} e^t & 0 \\ 0 & e^t \end{pmatrix}, \quad e^{tA_2} = \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix},$$

$$e^{tA_3} = \begin{pmatrix} \cosh(t) & \sinh(t) \\ \sinh(t) & \cosh(t) \end{pmatrix}, \quad e^{tA_4} = \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix}. \quad (4.99)$$

These forms seem more relevant to the study of physical problems than those given by the simple basis used in the previous section.

The commutator table with entries $[A_i, A_j]$ is given by

$A_i \backslash A_j$	A_1	A_2	A_3	A_4
A_1	0	0	0	0
A_2	0	0	$2A_4$	$2A_3$
A_3	0	$-2A_4$	0	$-2A_2$
A_4	0	$-2A_3$	$2A_2$	0

(4.100)

which is even simpler. Next we need a table of the adjoint actions $e^{A_i t} A_j e^{-A_i t}$ of the Lie group on the Lie algebra:

$A_i \backslash A_j$	A_1	A_2	A_3	A_4
A_1	A_1	A_2	A_3	A_4
A_2	A_1	A_2	$\cosh(2t) A_3 + \sinh(2t) A_4$	$\cosh(2t) A_4 + \sinh(2t) A_3$
A_3	A_1	$\cosh(2t) A_2 - \sinh(2t) A_4$	A_3	$\cosh(2t) A_4 - \sinh(2t) A_2$
A_4	A_1	$\cos(2t) A_2 - \sin(2t) A_3$	$\cos(2t) A_3 + \sin(2t) A_2$	A_4

(4.101)

As before this was computed using the series definition of the adjoint action and then identifying the series in the resulting formulas.

We now illustrate how to use an ordinary differential equation trick to evaluate one of the previous adjoint actions. If

$$X(t) = e^{A_2 t} A_3 e^{-A_2 t} = e^{[A_2, \cdot] t} A_3, \quad (4.102)$$

then

$$X'(t) = [A_2, X(t)], \quad X(0) = A_3. \quad (4.103)$$

If we set

$$X(t) = \begin{pmatrix} r(t) & s(t) \\ u(t) & v(t) \end{pmatrix} \quad (4.104)$$

then we need

$$\begin{pmatrix} r'(t) & s'(t) \\ u'(t) & v'(t) \end{pmatrix} = \left[\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} r(t) & s(t) \\ u(t) & v(t) \end{pmatrix} \right] = \begin{pmatrix} 0 & 2s(t) \\ 2u(t) & 0 \end{pmatrix}, \quad (4.105)$$

and

$$\begin{pmatrix} r(0) & s(0) \\ u(0) & v(0) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (4.106)$$

This gives

$$X(t) = \begin{pmatrix} 0 & e^{2t} \\ e^{-2t} & 0 \end{pmatrix} = e^{2t} \frac{A_3 + A_4}{2} + e^{-2t} \frac{A_3 - A_4}{2} = \cosh(2t) A_3 + \sinh(2t) A_4. \quad (4.107)$$

Of course, this checks with the power-series solution.

The Green's function of the the differential equation is represented in the form (3.64):

$$G(t, \tau) = e^{g_1(t, \tau) A_1} e^{g_2(t, \tau) A_2} e^{g_3(t, \tau) A_3} e^{g_4(t, \tau) A_4}, \quad (4.108)$$

where the g_i satisfy the initial conditions (3.65). Then the formulas for the exponentials of the basis elements (4.99) give the Green's function as a product of four matrices:

$$G = \begin{pmatrix} e^{g_1} & 0 \\ 0 & e^{g_1} \end{pmatrix} \begin{pmatrix} e^{g_2} & 0 \\ 0 & e^{-g_2} \end{pmatrix} \begin{pmatrix} \cosh(g_3) & \sinh(g_3) \\ \sinh(g_3) & \cosh(g_3) \end{pmatrix} \begin{pmatrix} \cos(g_4) & \sin(g_4) \\ -\sin(g_4) & \cos(g_4) \end{pmatrix}. \quad (4.109)$$

The determining equations (3.69) with $A(t)$ given by (4.80) have the form

$$\begin{aligned} a &= g'_1 + g'_2 - \sinh(2g_3) g'_4, \\ b &= \cosh(2g_2) g'_3 + \sinh(2g_2) g'_3 + \cosh(2g_2) \cosh(2g_3) g'_4 + \cosh(2g_3) \sinh(2g_2) g'_4, \\ c &= \cosh(2g_2) g'_3 - \sinh(2g_2) g'_3 - \cosh(2g_2) \cosh(2g_3) g'_4 + \cosh(2g_3) \sinh(2g_2) g'_4, \\ d &= g'_1 - g'_2 + \sinh(2g_3) g'_4. \end{aligned} \quad (4.110)$$

We solve equations (4.110) for g'_i to obtain the following nonlinear system of differential equations for the g_i :

$$\begin{aligned} g'_1 &= \frac{a+d}{2}, \\ g'_2 &= \frac{a-d}{2} + \left(\frac{b-c}{2} \cosh(2g_2) - \frac{b+c}{2} \sinh(2g_2) \right) \tanh(2g_3), \\ g'_3 &= \frac{b+c}{2} \cosh(2g_2) - \frac{b-c}{2} \sinh(2g_2), \\ g'_4 &= \left(\frac{b-c}{2} \cosh(2g_2) - \frac{b+c}{2} \sinh(2g_2) \right) \operatorname{sech}(2g_3). \end{aligned} \quad (4.111)$$

As before, even when $A(t)$ is constant these equations are seriously nonlinear and difficult to solve. But again, we can create many solvable examples by choosing the g_i functions appropriately. Again, for an example, we choose

$$g_1(t) = \alpha t, \quad g_2(t) = \beta t, \quad g_3(t) = \gamma t, \quad g_4(t) = \delta t, \quad (4.112)$$

and then the determining equations become

$$\begin{aligned} a(t) &= \alpha + \beta - \delta \sinh(2\gamma t), \\ b(t) &= \gamma \cosh(2\beta t) + \delta \cosh(2\beta t) \cosh(2\gamma t) + \gamma \sinh(2\beta t) + \delta \cosh(2\gamma t) \sinh(2\beta t), \\ c(t) &= \gamma \cosh(2\beta t) - \delta \cosh(2\beta t) \cosh(2\gamma t) - \gamma \sinh(2\beta t) + \delta \cosh(2\gamma t) \sinh(2\beta t), \\ d(t) &= \alpha - \beta + \delta \sinh(2\gamma t). \end{aligned} \quad (4.113)$$

If we assume the ordinary differential equation doesn't have a forcing term, then we only need $G(t) = G(t, 0)$ which is given by (4.109):

$$G(t) = \begin{pmatrix} e^{\alpha t} & 0 \\ 0 & e^{\alpha t} \end{pmatrix} \begin{pmatrix} e^{\beta t} & 0 \\ 0 & e^{-\beta t} \end{pmatrix} \begin{pmatrix} \cosh(\gamma t) & \sinh(\gamma t) \\ \sinh(\gamma t) & \cosh(\gamma t) \end{pmatrix} \begin{pmatrix} \cos(\delta t) & \sin(\delta t) \\ -\sin(\delta t) & \cos(\delta t) \end{pmatrix}, \quad (4.114)$$

or

$$\begin{aligned} g_{1,1} &= e^{\alpha t + \beta t} (\cos(\delta t) \cosh(\gamma t) - \sin(\delta t) \sinh(\gamma t)), \\ g_{1,2} &= e^{\alpha t + \beta t} (\cosh(\gamma t) \sin(\delta t) + \cos(\delta t) \sinh(\gamma t)), \\ g_{2,1} &= -e^{\alpha t - \beta t} (\cosh(\gamma t) \sin(\delta t) - \cos(\delta t) \sinh(\gamma t)), \\ g_{2,2} &= e^{\alpha t - \beta t} (\cos(\delta t) \cosh(\gamma t) + \sin(\delta t) \sinh(\gamma t)). \end{aligned} \quad (4.115)$$

In the applications that interest us, it is common for $A(t)$ to have trace zero, that is, $d = -a$, so that $g'_1 = 0$ and then $g_1(t, \tau) \equiv 0$. Another way of saying the same thing is that A_2 , A_3 and A_4 form a basis for the traceless matrices. In any case, this doesn't make much of a simplification.

Another important case is where $A(t)$ is skew symmetric, that is, $a = d = 0$ and $c = -b$:

$$\begin{aligned} g'_1 &= 0, \\ g'_2 &= b \cosh(2g_2) \tanh(2g_3), \\ g'_3 &= -b \sinh(2g_2), \\ g'_4 &= b \cosh(2g_2) \operatorname{sech}(2g_3). \end{aligned} \quad (4.116)$$

Although this is an impressive system of equations, its solution can easily be seen to be

$$g_1(t, \tau) = 0, \quad g_2(t, \tau) = 0, \quad g_3(t, \tau) = 0, \quad g_4(t, \tau) = \int_{\tau}^t b(s) ds. \quad (4.117)$$

This result is far easier to see starting from the fact that the Hermetian 2 by 2 matrices form a one dimensional space with A_4 being a basis.

4.2.3 The Pauli Spin Basis

We now look at 2 by 2 skew-symmetric matrices and choose the Pauli Spin matrices for a basis:

$$A_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & -j \\ j & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (4.118)$$

Then any skew-symmetric matrix can be represent by

$$A(t) = a(t) A_1 + b(t) A_2 + c(t) A_3 = \begin{pmatrix} c(t) & a(t) - j b(t) \\ a(t) + j b(t) & -c(t) \end{pmatrix} \quad (4.119)$$

The exponentials of the basis elements are

$$\begin{aligned} e^{t A_1} &= \begin{pmatrix} \cosh(t) & \sinh(t) \\ \sinh(t) & \cosh(t) \end{pmatrix}, \\ e^{t A_2} &= \begin{pmatrix} \cosh(t) & -j \sinh(t) \\ j \sinh(t) & \cosh(t) \end{pmatrix}, \\ e^{t A_3} &= \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}. \end{aligned} \quad (4.120)$$

The commutator table with entries $[A_i, A_j]$ is given by

$A_i \backslash A_j$	A_1	A_2	A_3
A_1	0	$2j A_3$	$-2j A_2$
A_2	$-2j A_3$	0	$2j A_1$
A_3	$2j A_2$	$-2j A_1$	0

(4.121)

Next we need a table of the adjoint actions $e^{A_i t} A_j e^{-A_i t}$ of the Lie group on the Lie algebra:

$A_i \backslash A_j$	A_1	A_2	A_3
A_1	A_1	$A_2 \cosh(2t) + j A_3 \sinh(2t)$	$A_3 \cosh(2t) - j A_2 \sinh(2t)$
A_2	$A_1 \cosh(2t) - j A_3 \sinh(2t)$	A_2	$A_3 \cosh(2t) + j A_1 \sinh(2t)$
A_3	$A_1 \cosh(2t) + j A_2 \sinh(2t)$	$A_2 \cosh(2t) - j A_1 \sinh(2t)$	A_3

(4.122)

As before this was computed using the series definition of the adjoint action and then identifying the series in the resulting formulas.

The Green's function of the the differential equation is represented in the form (3.64):

$$G(t, \tau) = e^{g_1(t, \tau) A_1} e^{g_2(t, \tau) A_2} e^{g_3(t, \tau) A_3}, \quad (4.123)$$

where the g_i satisfy the initial conditions (3.65). Then the formulas for the exponentials of the basis elements (4.120) give the Green's function as a product of three matrices:

$$G = \begin{pmatrix} \cosh(g_1) & \sinh(g_1) \\ \sinh(g_1) & \cosh(g_1) \end{pmatrix} \begin{pmatrix} \cosh(g_2) & -j \sinh(g_2) \\ j \sinh(g_2) & \cosh(g_2) \end{pmatrix} \begin{pmatrix} e^{g_3} & 0 \\ 0 & e^{-g_3} \end{pmatrix}. \quad (4.124)$$

The determining equations (3.69) with $A(t)$ given by (4.119) have the form

$$\begin{aligned} a(t) &= g'_1 + j \sinh(2 g_2(t)) g'_3, \\ b(t) &= \cosh(2 g_1(t)) g'_2 - j \cosh(2 g_2(t)) \sinh(2 g_1(t)) g'_3, \\ c(t) &= j \sinh(2 g_1(t)) g'_2 + \cosh(2 g_1(t)) \cosh(2 g_2(t)) g'_3. \end{aligned} \quad (4.125)$$

We solve equations (4.125) for g'_i to obtain the following nonlinear system of differential equations for the g_i :

$$\begin{aligned} g'_1(t) &= a \cosh(2 g_1(t))^2 - j c \cosh(2 g_1(t)) \tanh(2 g_2(t)) \\ &\quad - \sinh(2 g_1(t)) (a \sinh(2 g_1(t)) + b \tanh(2 g_2(t))) \\ g'_2(t) &= b \cosh(2 g_1(t)) + j c \sinh(2 g_1(t)) \\ g'_3(t) &= \operatorname{sech}(2 g_2(t)) (c \cosh(2 g_1(t)) - j b \sinh(2 g_1(t))) \end{aligned} \quad (4.126)$$

4.3 Three-Dimensional Example

The full three-dimensional case involves nine parameters and thus there isn't much use in tabulating all of this. However, the skew-symmetric 3 by 3 matrices are only three dimensional. A convenient basis is

$$A_1 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}. \quad (4.127)$$

Then any skew-symmetric matrix $A(t)$ can be represented as

$$A(t) = a(t) A_1 + b(t) A_2 + c(t) A_3 = \begin{pmatrix} 0 & a(t) & c(t) \\ -a(t) & 0 & b(t) \\ -c(t) & -b(t) & 0 \end{pmatrix}. \quad (4.128)$$

The exponentials of the basis elements are

$$\begin{aligned} e^{t A_1} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(t) & \sin(t) \\ 0 & -\sin(t) & \cos(t) \end{pmatrix}, \\ e^{t A_2} &= \begin{pmatrix} \cos(t) & 0 & -\sin(t) \\ 0 & 1 & 0 \\ \sin(t) & 0 & \cos(t) \end{pmatrix}, \\ e^{t A_3} &= \begin{pmatrix} \cos(t) & \sin(t) & 0 \\ -\sin(t) & \cos(t) & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned} \quad (4.129)$$

These matrices describe a rotation about each of the axes.

The commutator table with entries $[A_i, A_j]$ is given by

$A_i \backslash A_j$	A_1	A_2	A_3
A_1	0	A_3	$-A_2$
A_2	$-A_3$	0	A_1
A_3	A_2	$-A_1$	0

(4.130)

which is well known for the rotations. The table of the adjoint actions $e^{A_i t} A_j e^{-A_i t}$ of the Lie group on the Lie algebra are:

$A_i \backslash A_j$	A_1	A_2	A_3
A_1	A_1	$A_2 \cos(t) + A_3 \sin(t)$	$A_3 \cos(t) - A_2 \sin(t)$
A_2	$A_1 \cos(t) - A_3 \sin(t)$	A_2	$A_3 \cos(t) + A_1 \sin(t)$
A_3	$A_1 \cos(t) + A_2 \sin(t)$	$A_2 \cos(t) - A_1 \sin(t)$	A_3

(4.131)

The Green's function of the the differential equation is represented in the form (3.64):

$$G(t, \tau) = e^{g_1(t, \tau) A_1} e^{g_2(t, \tau) A_2} e^{g_3(t, \tau) A_3}, \quad (4.132)$$

where the g_i satisfy the initial conditions (3.65). Then the formulas for the exponentials of the basis elements (4.129) give the Green's function as a product of three matrices:

$$G(t, \tau) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(g_1(t, \tau)) & \sin(g_1(t, \tau)) \\ 0 & -\sin(g_1(t, \tau)) & \cos(g_1(t, \tau)) \end{pmatrix}$$

$$\begin{pmatrix} \cos(g_2) & 0 & -\sin(g_2) \\ 0 & 1 & 0 \\ \sin(g_2) & 0 & \cos(g_2) \end{pmatrix} \begin{pmatrix} \cos(g_3) & \sin(g_3) & 0 \\ -\sin(g_3) & \cos(g_3) & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (4.133)$$

The determining equations (3.69) can be evaluated using the adjoint action table to give

$$\begin{aligned} A(t) = & (g'_1 + \sin(g_2) * g'_3) A_1 \\ & + (\cos(g_1) * g'_2 - \cos(g_2) * \sin(g_1) * g'_3) A_2 \\ & + (\sin(g_1) * g'_2 + \cos(g_1) * \cos(g_2) * g'_3) A_3 \end{aligned} \quad (4.134)$$

With $A(t)$ given by (4.128), these equations can be solved for g'_i to obtain the following nonlinear system of differential equations for the g_i :

$$\begin{aligned} g'_1 &= a - \tan(g_2) (c \cos(g_1) + b \sin(g_1)), \\ g'_2 &= b \cos(g_1) + c \sin(g_1), \\ g'_3 &= \sec(g_2) (c \cos(g_1) - b \sin(g_1)). \end{aligned} \quad (4.135)$$

The nonlinear system of equations is not easy to solve, but as before, we can create a large family of solvable examples by choosing the g_i functions. If we look only at the initial value problem (no source terms), and if we choose the a , b , and c that determine $A(t)$ in (4.134) as

$$\begin{aligned} a(t) &= \alpha + \gamma \sin(\beta t), \\ b(t) &= \beta \cos(\alpha t) - \gamma \cos(\beta t) \sin(\alpha t), \\ c(t) &= \gamma \cos(\alpha t) \cos(\beta t) + \beta \sin(\alpha t), \end{aligned} \quad (4.136)$$

then the greens function $G(t) = G(t, 0)$ is given by the product of three simple matrices (4.133):

$$\begin{aligned} G(t) = & \begin{pmatrix} \cos(\alpha t) & \sin(\alpha t) & 0 \\ -\sin(\alpha t) & \cos(\alpha t) & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ & \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\beta t) & \sin(\beta t) \\ 0 & -\sin(\beta t) & \cos(\beta t) \end{pmatrix} \\ & \begin{pmatrix} \cos(\gamma t) & 0 & \sin(\gamma t) \\ 0 & 1 & 0 \\ -\sin(\gamma t) & 0 & \cos(\gamma t) \end{pmatrix}. \end{aligned} \quad (4.137)$$

So, in summary, we can use Lie theory to convert systems of ordinary differential equations with variable coefficients into systems of nonlinear equations that are difficult to solve. However, the resulting formulas allow us to create a large number of solvable examples (along with the solutions).

5 Conclusions

We have laid the foundations of the theoretical results necessary to apply Lie-algebraic methods to the solution of important problems in transmission-line and wave-launcher problems. Future work will involve an integration of the current techniques with the theory of matrizants- propagators-Green's functions and then applications to non-trivial examples.

References

- [1] C.E. Baum and H.N. Kritikos, *Electromagnetic Symmetry* (in the Electromagnetics Library series), Taylor and Francis, 1995.
- [2] C.E. Baum, *Coupled Transmission-Line Model of Periodic Array of Wave Launchers*, Sensor and Simulation Notes 313, Air Force Weapons Laboratory, December 1988.
- [3] C.E. Baum, *Canonical Examples for High-Frequency Propagation on Unit Cell of Wave-Launcher Array*, Sensor and Simulation Notes 317, Air Force Weapons Laboratory, April 1989.
- [4] C.E. Baum, *High-Frequency Propagation on Nonuniform Multiconductor Transmission Lines in Uniform Media*, Interaction Notes 463, Air Force Weapons Laboratory, March 1988.
- [5] C.E. Baum, *Nonuniform Multiconductor Transmission Lines*, Interaction Notes 516, Air Force Phillips Laboratory, February 1986.
- [6] J.D. Dollard and C.N. Friedman, *Product Integration*, Encyclopedia of Mathematics and Its Applications, (F.E. Browder, ed.), Addison-Wesley Publishing Company, 1979.
- [7] F.R. Gantmacher, *The Theory of Matrices*, Chelsea, 1959.
- [8] D.V. Giri, *A Family of Canonical Examples for High Frequency Propagation on Unit Cell of Wave-Launcher Array*, Sensor and Simulation Notes 318, Air Force Weapons Laboratory, June 1989.
- [9] L. Hlavaty, B. Wolf and S. Steinberg, *Riccati Equations and Lie Series*, Journal of Mathematical Analysis and Applications, **104** (1984), 246-263.
- [10] S. Steinberg, *Lie Series, Lie Transformations and Their Applications*, Lie Methods in Optics, (J. Sánchez Mondragón and K. B. Wolf, eds.), Lecture Notes in Physics, Springer Verlag, 1985, 45-103.
- [11] S. Steinberg, *Factored Product Expansions of Solutions of Nonlinear Differential Equations*, SIAM journal on Mathematical Analysis, **15** (1984), 108-115.
- [12] S. Steinberg, *Lie Series and Nonlinear Ordinary Equations*, Journal of Mathematical Analysis and Applications, **101** (1984), 39-63.
- [13] S. Steinberg, *Applications of the Lie Algebraic Formulas of Baker, Campbell, Hausdorff and Zassenhaus to the Explicit Solutions of Partial Differential Equations*, Journal of Differential Equations, **26** (1977), 404-434.
- [14] B.G. Wybourne, *Classical Groups for Physicists* Wiley, 1974.